Emigration by Educational Attainment and Growth: Cross-Country Evidence and Growth Implications of Immigration: Evidence from U.S. Industries

Author: Shoghik Hovhannisyan

Boston College Electronic Thesis or Dissertation, 2014

Boston College

The Graduate School of Arts and Sciences

Department of Economics

Emigration by Educational Attainment and Growth: Cross-Country Evidence and Growth Implications of Immigration: Evidence from U.S. Industries

a dissertation

by

SHOGHIK HOVHANNISYAN

submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

December, 2014

Abstract

Emigration by Educational Attainment and Growth: Cross-Country Evidence and

Growth Implications of Immigration: Evidence from U.S. Industries

by

SHOGHIK HOVHANNISYAN

Dissertation Committee:

Professor FABIO SCHIANTARELLI (Chair)

Professor SCOTT FULFORD

Professor CHRISTOPHER BAUM

This thesis includes two essays that analyze growth implications of emigration and immigration. The first chapter studies the impact of emigrants with different education levels on their home countries' GDP per worker and its factors obtained by a production function decomposition. It uses migration data from 195 countries of origin to 30 major destination OECD countries in 1990 and 2000 and applies an instrumental variable approach to correct for endogeneity bias in estimating this impact. Pull factors of migration such as demand for emigrants' labor in destination countries and migrants' networks serve as a basis for instrument construction. Estimation results indicate that growth in emigration rates increases growth in GDP per worker in low and lower-middle income countries for all education groups of emigrants, primarily driven by improvements in total factor productivity (TFP). In contrast, there is no robust significant impact of emigration on other components of GDP.

The second chapter studies the impact of immigrant labor on GDP per worker in the U.S. and its components obtained by a production function decomposition, including total factor productivity (TFP), the capital-output ratio, average hours worked, and skill intensity, defined as a productivity-weighted Constant Elasticity of Substitution (CES) function of high-skill and low-skill workers. It uses industry-level data over the period of 1960–2005 and applies two-step Difference Generalized Method of Moments

(GMM) with instruments constructed using past distributions of immigrants across industries. The estimation results show that GDP per worker in an industry increases by about 2.24–2.63 percent in response to a one percent increase in the share of immigrants in total employment of the industry. These results are primarily driven by TFP growth with a magnitude of 2.08–2.21 and average hours worked: 0.23–0.29. However, these results are not robust to inclusion of the lagged dependent variables.

Contents

Chapter 1

Emigration by Educational Attainment and Growth: Cross-Country Evidence

1.1 Introduction

Emigration became an important factor in the development of migrant sending countries. The number of foreign-born individuals residing in 30 OECD countries increased from 42 million to nearly 59 million over the period of 1990–2000. This paper studies the impact of emigration for different education levels on GDP per worker and its factors obtained by a production function decomposition. It uses data on emigration from 195 countries of origin to 30 OECD destination countries which account for about 70 percent of total emigration and 90 percent of skilled emigration in the world.

Emigration can have both a negative and a positive impact on migrant-sending countries. On one hand, these countries face deprivation of their labor force especially the most educated, known in the literature as a brain drain phenomenon. On the other hand, they benefit from emigration in several ways. Migrants remit money home and these financial flows account for a significant share of GDP in many developing countries. Remittances relax financial constraints of households which have a member or relative abroad and can increase not only their consumption of goods and services but also expenditures on health and education, thus having both short-term and long-term effects on GDP. Emigration might also promote transfer of technology and knowledge across countries by facilitating more foreign direct investment (FDI), trade, and other partnerships through established diasporas abroad and their networks. Finally, the high probability of emigration of educated labor raises returns to education and, therefore, might lead to higher investments in education. As not everyone has a chance to migrate, this has the potential to increase overall human capital in the migrant-sending country.

There is a large empirical literature on emigration effects in migrant-sending countries and their various transmission mechanisms. Macroeconomic studies using cross-country data provide mixed evidence on the effects of emigration and remittances on growth and its drivers, with results highly dependent on the econometric approaches and instruments used to control for endogeneity bias. Estimations by Cartinescu et al. (2009) and Acosta et al. (2008) show an increase in growth due to remittances, while Chami et al. (2009) find a negative effect of remittances on growth volatility. However, Barajas et al. (2009) conclude that remittances have no impact on economic growth in their cross-country analysis. At the same time, Easterly and Nyarko (2009) find no significant impact of brain drain or outflow of high-skill migrants on GDP growth in African countries by using distance from France, U.K., and U.S., and population as instruments to address endogeneity issues. In addition, Gould (1994) shows that U.S. bilateral trade is larger with countries that send more migrants to the U.S. and Head et al. (1998) estimate similar effects for Canada.

A significant fraction of the emigration literature discusses its impact on human capital of migrant-sending countries. Papers by Beine et al. (2008), Docquier et al. (2008), Docquier et al. (2007), and Easterly and Nyarko (2009) use a variety of instruments such as total population size; migration stocks at the beginning of the period; geographical proximity to developed countries; dummy variables for small islands, landlocked, least-developed, and oil exporting countries; former colonial links; etc., to correct for endogeneity bias in estimating these effects. These papers find a positive, significant impact of emigration on human capital formation in countries of origin due to a higher propensity to migrate for more educated people, which increases investments in education.

Studies using household, firm, or individual-level data discuss various transmission mechanisms of emigration impact on growth. Yang (2008) finds an increase in remittances to households in the Philippines at the time of the Asian financial crisis, consistent with consumption smoothing. In contrast, government transfers have no impact on remittances in Mexico, according to Teruel and Davis (2000) or in Honduras and Nicaragua, according to Nielsen and Olinto (2007). Woodruff and Zenteno (2006) find that migration is associated with higher investment levels and profits when analyzing data on self-employed workers and small firm owners in urban areas of Mexico. Also, using panel data from rural Pakistan, Adams (1998) shows that availability of remittances helps increase investment in rural assets by raising the marginal propensity to invest for migrant households. In addition, remittances increase households' school attendance in El Salvador according to Cox and Ureta (2003) and improve health outcomes in Mexico according to Hilderbrandt and McKenzie (2005). Finally, a study by Saxenian (2002) concludes that emigration of India's high-skill labor to Silicon Valley increased trade with and investment from the U.S., promoting creation of local high-technology industries. In terms of labor market outcomes, Mishra (2007), Aydemir and Borjas (2007), and Hanson (2007) find a positive correlation between wages and emigration in Mexico.

This paper contributes to the emigration literature by studying the growth implications of emigration across different education groups of population using a new econometric approach. There is an extensive literature on immigration impact on the growth of migrant hosting countries. However, the emigration–growth or brain drain–growth literature is highly limited due to the lack of the time series data on emigration, especially for different education groups and the absence of good instruments to address potential reverse causality and endogeneity issues. Labor emigration might have an impact on GDP growth, as it directly affects the labor or human capital in the migrant sending countries. At the same time, low levels of GDP might create economic incentives for people to migrate to higher income countries with better welfare. Also, there might be other factors affecting both the GDP and emigration such as natural disasters, wars, etc. The only paper which studies the impact of emigration and in, particular, the effect of brain drain on growth is Easterly and Nyarko (2009) which uses distance from France, U.K., and U.S.; dummy variables for former colonies of Great Britain and France; and population as instruments to address potential reverse causality and endogeneity issues. However, the study also acknowledges the limitations of the instruments used in the estimates, which perform poorly on both weak-instruments and overidentifying-restrictions tests.

To address endogeneity and simultaneity bias in Ordinary Least Squares (OLS) estimation, this paper applies an Instrumental Variable (IV) approach. A new instrument is constructed based on pull factors of migration and migrants' networks to estimate how changes in population due to emigration affect the growth of GDP per capita and its components. An increase in total immigration stocks of destination countries is primarily driven by either changes in immigration policies or labor demand, and is taken as exogenous to developments in countries of origin. At the same time, there is a strong tendency for immigrant groups to go where previous immigrants from their country have already migrated (Card, 2001). An increase in demand for migrants with a particular education in a destination country will therefore create a pull for migrants that varies across sending countries.

Estimation results indicate that emigration rates increase the GDP per worker in low and lower-middle income countries in all specifications. These results hold for all education groups of emigrants. All coefficients across different econometric specifications for total number of migrants are significant, within the range of $1.7 - 2.29$, and robust to the inclusion of different control variables. The magnitude of the coefficients for secondary and tertiary educated emigrants is lower (0.5) than for the total number of emigrants, possibly reflecting the importance of education independent channels. The coefficients range from 0.04 to 0.10 for the impact of the change in the number of emigrants with tertiary education relative to their respective population, and, thus, have smaller values than those for secondary and tertiary educated emigrants and for the total number of emigrants. However, a test for a difference in the estimated coefficients between different country groups reveals that only coefficients of the total emigration–population change in low and lower-middle income countries are significantly different from others.

Among different components of GDP per worker, TFP is a major source for improvement in low and lower-middle income countries, but the education level is not the underlying reason for the increase in productivity. These changes in TFP might be a result of trade or other cross-country partnerships facilitated by established diasporas abroad which lead to a transfer of knowledge and technology.

The rest of the paper is organized as follows. Section 2 introduces the theoretical and econometric framework for growth accounting. Section 3 describes the data and construction of the variables. Section 4 summarizes the statistics and trends of the key variables. Section 5 discusses the choice of instruments. The estimation results are presented in Section 6. Finally, Section 7 concludes.

1.2 Theoretical and Econometric Framework

The empirical strategy estimates the impact of different education groups of emigrants on the GDP per worker. The reduced form equation (1.1) is estimated using Ordinary Least Squares (OLS) and the Instrumental Variable (IV) Approach to address endogeneity problems.

$$\ln y_{it} = \alpha_i^k + \lambda_t^k + \beta^k \frac{E_{it}^k}{P_{i,t}^k} + \epsilon_{it}^k \tag{1.1}$$

where $\ln y_{it}$ is the logarithm of GDP per worker in country i in period t, $\frac{E_{it}^k}{P_{i,t}^k}$ is an emigration-population ratio for education group k in country i in period t, and ϵ_{it}^k is a zero-mean random shock.

Time and country specific effects are captured by λ_t^k and α_i^k, the latter, as a source of endogeneity, can be eliminated by taking the first difference of the right and left-hand side variables as presented in Equation (1.2) below:

$$\Delta \ln y_{it} = \Delta\lambda_t^k + \beta^k \Delta \frac{E_{it}^k}{P_{i,t}^k} + \Delta\epsilon_{it}^k \tag{1.2}$$

Equation (1.2) is modified further for estimation purposes given data constraints and use of some approximations. The first variable on the right-hand side of Equation (1.2), $\Delta\lambda_t^k$, becomes a constant as there is only one period available given a lack of migration data over an extended period of time. Also, considering an insignificant change in population over a decade, the paper uses $\frac{\Delta E_{it}^k}{P_{i,t-1}^k}$ instead of $\Delta\frac{E_{it}^k}{P_{i,t}^k}$. Finally, the error term $\Delta\epsilon_{it}^k$ becomes a Moving Average (MA(1)) process which is important to consider in the construction of the instrument in the IV approach.

One might also be interested in the impact of emigration on the components of GDP per worker growth and this paper uses a growth accounting framework to analyze these effects. To study the channels of emigration impact it decomposes the GDP into three factors using the following Cobb–Douglas production function as in Caselli (2005):

$$Y_{it} = A_{it} K_{it}^{\alpha} (L_{it} h_{it})^{1-\alpha} \tag{1.3}$$

where A_{it} is TFP, K_{it} is the aggregate capital stock, α is the capital share in GDP, and $(L_{it} h_{it})$ is a quality adjusted workforce, with the number of workers L_{it} multiplied by their average human capital h_{it}, in country i and period t. In per-worker terms the production function can be written as:

$$y_{it} = A_{it} k_{it}^{\alpha} (h_{it})^{1-\alpha} \tag{1.4}$$

where k_{it} is the capital-labor ratio (K_{it}/L_{it}). This equation can be rewritten as a growth of GDP per worker by taking the logarithm of each side and their derivatives:

$$\hat{y}_{it} = \hat{A}_{it} + \alpha \hat{k}_{it} + (1 - \alpha)\hat{h}_{it} \tag{1.5}$$

where hat indicates a derivative, approximated by the log change in discrete time.

Physical capital, K_{it}, is constructed using the perpetual inventory method:

$$K_{it} = I_{it} + (1 - \delta)K_{it-1} \tag{1.6}$$

where I_{it} is investment in country i and period t and δ is a depreciation rate. The initial capital stock K_{it}^0 is obtained from the steady-state expression for the capital stock in the Solow model:

$$K_{it}^0 = \frac{I_i^0}{g_i + \delta} \tag{1.7}$$

where I_i^0 is a value of the investment series in the first available year and g_i is an average geometric growth rate for the investment series between the first available year and 2000 for country i. To compute the time series for K_{it}, investment in respective years is added to the initial capital stock.

The average human capital h_{it} is a function of average years of schooling in the population as expressed in the following equation:

$$h_{it} = e^{\phi(s_{it})} \tag{1.8}$$

where s_{it} is average years of schooling in country i and period t and $\phi(s_{it})$ is a piecewise linear function with slope 0.13 for $s_{it} \leq 4$, 0.10 for $4 < s_{it} \leq 8$, and 0.07 for $8 < s_{it}$ (Hall and Jones, 1999). This function resembles the log-linear functional relationship between wages and years of education in the Mincerian approach, where wages are assumed to be proportional to human capital given the production function and perfect competition. Since international data on education and wages suggest that there are some differences in marginal rates of return across countries, those differences are introduced with the convexity. Finally, TFP, A_{it}, is constructed as a residual.

1.3 Data Description

This study uses the migration dataset by Docquier, Lowell, and Marfouk (2008) which provides the number of migrants from 195 migrant-sending countries to 30 main destination OECD countries in 1990 and 2000. These emigration stocks account for about 70 percent of total emigration and 90 percent of skilled emigration in the world. The dataset classifies emigrants into three groups based on education: high-skill, medium-skill, and low-skill emigrants with respectively a post-secondary, an upper secondary, and a primary or no formal education. It also provides emigration rates for each education group defined as a share of emigrants in the total native population including residents and emigrants in the same education category. Data on migration in 1980 for constructing an instrument are taken from the Global Bilateral Migration Database of the World Bank.

Country-level aggregate variables including the employment-population ratio, GDP per worker, capital per worker, and labor inputs are obtained from the Penn World Tables (PWT) by Heston, Summers and Bettina (PWT 7.0). First, the number of workers in each country i and year t is computed as $(rgdpch_{it} * pop_{it}/rgdpwok_{it})$, where $rgdpch_{it}$ is a PPP converted GDP per capita (Chain Series) at 2005 constant prices, pop_{it} is a population, and $rgdpwok_{it}$ is a PPP Converted GDP Chain per worker at 2005 constant prices. The capital-worker ratio k is computed using the perpetual inventory method:

$$K_{it} = I_{it} + (1 - \delta)K_{it-1} \tag{1.9}$$

where I_{it} is investment and δ is a depreciation rate. I_{it} is computed as $(rgdpl_{it} * pop_{it} * ki_{it})$, where $rgdpl_{it}$ is a PPP converted GDP per capita (Laspeyres) measure at 2005 constant prices, pop_{it} is population, and ki_{it} is the investment share of PPP converted

GDP per capita at 2005 constant prices in country i and year t. The depreciation rate δ equals 0.06, which is a conventional value used in the literature (see for example Caselli, 2005). In addition, PWT 7.0 provides data on several control variables discussed below such as government size (kg_{it}) and openness of the economy ($openk_{it}$) measured respectively as the shares of government expenditures and trade, including exports and imports, in GDP.

The average human capital h_{it} is constructed using average years of schooling in the population over 25 years old from the Barro–Lee dataset (Barro and Lee, 2000). As in Docquier and Marfouk (2006), human capital indicators are replaced with those from De La Fuente and Domenech (2002) for OECD countries. For countries where Barro and Lee measures are missing, the proportion of educated individuals is predicted using the Cohen and Soto (2007) measures. In the result, there are 25 missing observations for 1990 and 35 for 2000, accounting respectively for 15 and 20 percent of total observations. These missing observations are imputed using GDP per worker: the coefficients of the regression of average years of schooling on the GDP per worker are used to compute the missing observations. Finally, TFP is constructed as a residual.

1.4 Summary Statistics and Trends

The median and mean growth rates, by decade, of all the dependent variables used in estimation: GDP per worker, capital per worker, human capital, and TFP, are reported in Table $B.2$. These data are also disaggregated by country groups used in the estimates, which include 115 non-high income countries and 74 low and lower-middle income country groups (Table $B.1$).

There was a huge variation in the median and mean growth rates of GDP per worker for different country groups. The median and mean growth rates of GDP per worker accounted respectively for 0.11 and 0.16 for all countries over the period of 1990–2000. Despite a low level of GDP per worker in less-developed countries and catching-up effects between low and high income countries, the median and mean growth rates of non-high income countries were much lower: respectively, 0.04 and 0.07. Moreover, the low and lower-middle income countries demonstrated a negative median growth rate with a magnitude of -0.02 and a mean positive growth rate of 0.02.

Median and mean growth rates of capital per worker demonstrated similar variations across various country groups as shown in Table $B.2$. Due to low level of savings and imperfect capital flows from advanced economies to developing countries and high capital intensity in the latter group, the median and mean growth rates of the capital per worker were less in the non-high income countries: 0.03 and 0.06, and in the low and lower-middle income countries: 0.02 and 0.03 as opposed to 0.04 and 0.11 for all countries.

Table $B.2$ also shows that there were small differences in human capital across various country groups with median and mean growth rates in the range of $0.07 - 0.09$. The mean and median growth rates of TFP over 1990–2000 had negative magnitudes across all country groups. As in the case of GDP per worker and capital per worker, the mean and median growth rates of TFP in all countries were higher than in other country groups.

The emigration specific variables' trends in Table $B.3$ show that median and mean changes in emigration relative to population and the emigration-population ratio. These variables are reported for three education categories: all emigrants; emigrants with secondary and tertiary education; and emigrants with tertiary education. In addition, these variables are disaggregated by country income groups.

1.5 Choice of Instruments

This study estimates the impact of emigration on GDP per worker and its factors obtained by a production function decomposition in migrant-sending countries using cross-country data over the period of 1990–2000. It analyzes the impact of emigration for three different education groups of the native population: for all levels of education, those with secondary and tertiary education, and those with tertiary education. Distinguishing across these groups is important in understanding to what extent education of emigrants matters for development of their home countries. Low-skill emigration can simply lead to a decline in labor or influence countries of origin through remittances, promotion of FDI and trade, etc. In addition to these channels, high-skill emigration directly reduces the level of human capital in the migrant-sending countries but might contribute to investments in education, given a higher likelihood to emigrate for individuals with more education, as emphasized in the literature.

Estimating the impact of contemporaneous emigration on the growth of migrant sending countries in the reduced form equations by Ordinary Least Squares (OLS) might generate bias in the coefficients due to reverse causality or endogeneity. For example, emigration of highly educated people might decrease GDP per worker in the sending countries, given a higher marginal productivity of high-skill labor compared to low-skill labor. At the same time, a low level of GDP per worker might induce migration of more people, both high-skilled and low-skilled, to higher-income countries with better standards of living. In terms of endogeneity, there might be other factors driving both emigration and GDP per worker such as civil wars, weak institutions, natural disasters, etc., which might reduce GDP growth and increase emigration to countries with better opportunities. To address these econometric problems the paper applies an IV approach and constructs instruments that are uncorrelated with migrant sending countries' specific shocks but are correlated with emigration. In constructing the instruments it is important to consider that the estimated Equation 1.2 is expressed in terms of first-differences. This means that the error term follows a Moving Average (MA(1)) process: $\Delta\epsilon_{it}^k = \epsilon_{it}^k - \epsilon_{i,t-1}^k$. Therefore, any instrument which uses the first lag of respective variables would be correlated with the error term.

Migrants' networks and pull factors of migration provide variations in emigration exogenous to migrant-sending countries' conditions and, therefore, can serve as a basis for constructing an instrument. There are economic incentives for labor mobility between OECD countries and the rest of the world given a huge gap in income levels. In these circumstances, migrants' networks stimulate migration flows, as having individuals from the same countries of origin provides access to jobs and other information, substantially reducing migration costs. Figure $B.1$ depicts these network externalities, indicating that countries with high emigration rates or with large diasporas in 1990 tend to have high emigration rates in 2000 as well. Each point on these graphs shows a share of emigrants in the total native population in each migrant-sending country in 1990 and 2000 for three education levels: all, secondary and tertiary, and tertiary, thus highlighting the key role of networks in a choice to emigrate.

Figure $B.2$ demonstrates that countries with large stock of emigrants tend to have sizable growth in the number of emigrants across all education groups. Each point on this graphs indicates a migrant-sending country. In addition, Figure $B.3$ and Figure $B.4$ capture the network effects through the specific country examples. Figure $B.3$ focuses on the USA to demonstrate that countries with a high share of migrants in the total number of immigrants in the USA in 1990 tend to have a high share of migrants in 2000 as well. Finally, Figure $B.4$ illustrates the network effects for total emigration from India and Philippines, where the distribution of migrants across major destination countries remains relatively stable over time. The literature mostly discusses networks as a decisive factor in migrants' location choices in the context of subnational data. The methodology proposed by Card (2001) uses previous settlements of immigrants as an instrument in studying labor market effects of immigration across geographical regions in the U.S. This study expands the existing literature by using network effects for country level analysis. The growth in the total number of immigrants in each of 30 destination countries, which might be a combination of different factors such as increases in overall labor demand and changes in immigration policies, is also used to construct the instrument. Assuming there are economic incentives for emigration from developing to developed countries, a higher demand from destination countries triggers more emigration. At the same time, as migrants' networks or diasporas play an important role in migrants' destination choices, an increase in the number of immigrants in the destination country from different countries of origin is likely to be proportional to the sizes of their diasporas.

The independent variable in equation (1.1) can be rewritten in the following way:

$$\frac{E^k_{i,2000} - E^k_{i,1990}}{P^k_{i,1990}} = \frac{E^k_{i,1990}(1 + G^k_{i,2000-1990}) - E^k_{i,1990}}{P^k_{i,1990}} = \frac{E^k_{i,1990} G^k_{i,2000-1990}}{P^k_{i,1990}} = \frac{\sum_j E^k_{ij,1990} G^k_{\hat{i}j,2000-1990}}{P^k_{i,1990}} \tag{1.10}$$

where $E^k_{i,2000}$ and $E^k_{i,1990}$ are respectively the actual numbers of emigrants in country i with education level k in 2000 and 1990, $P^k_{i,1990}$ is population in country i in 1990 with education level k, and G^k_i is the growth rate in the number of emigrants with education level k in country i in 2000 relative to 1990.

The IV approach consists of the following steps. First, the growth in the total number of immigrants in 2000 relative to 1990 is computed for each of 30 OECD destination countries and for each education group using the actual number of immigrants:

$$G^k_{\hat{i}j,2000-1990} = \frac{E^k_{\hat{i}j,2000} - E^k_{\hat{i}j,1990}}{E^k_{\hat{i}j,1990}} \tag{1.11}$$

where $G^k_{\hat{i}j,2000-1990}$ is the growth rate in the total number of immigrants with education level k in destination country j in 2000 relative to 1990 to be used for country of origin i, $E^k_{\hat{i}j,2000}$ and $E^k_{\hat{i}j,1990}$ are respectively the actual numbers of immigrants in country j with education level k excluding immigrants from country i in 2000 and 1990. Excluding the number of migrants from country i in the total number of immigrants in destination countries eliminates any impact of country of origin i on the increase of immigration in destination countries. Therefore, this measurement of immigration growth in destination countries is purely demand driven, which ensures the exogeneity of the constructed instrument [1].

[1] The results are similar in terms of sign and significance when $E^k_{\hat{i}j,2000}$ and $E^k_{\hat{i}j,1990}$ include immigrants from country i as well.

Next, the paper suggests several instruments for different education groups of migrants. For all education groups in IV1 it applies constructed destination countries' growth rates to the number of migrants from each country of origin i in the respective destination country j in 1980 in order to impute the number of migrants in 2000. These imputed numbers of migrants are summed across destination countries to obtain the total number of emigrants for each migrant sending country i:

$$Z_{1i}^{k} = \frac{\sum_j E_{ij,1980} G_{\hat{i}j,2000-1990}^{k}}{P_{i,1980}} \tag{1.12}$$

This instrument meets the exclusion restriction as it correlates with emigration but not with other confounding factors in migrant-sending countries. The stock of emigrants in 1980 is exogenous to any shocks in migrant-sending countries in 1990 possibly captured in the error term of the reduced form equation (1.1). At the same time, it is highly correlated with the number of migrants in 1990 as the variable is a stock of migrants, reflecting the importance of diasporas in reducing migration costs and stimulating more emigration. Inflating the stock of emigrants in 1980 by the constructed growth rates of immigration in destination countries relies on the pull factors of emigration such as demand for migrants in destination countries excluding a given migrant sending country and migrants' networks, recognized as an essential factor in migrants' location choices.

Alternatively, for the total number of migrants in IV2, the paper uses growth rates of migrants both in 1990 relative to 1980 and 2000 relative to 1990 constructed as in Equation 1.9 to impute the total number of emigrants:

$$Z_{2i}^{1} = \frac{\sum_j E_{ij,1980}(1 + G_{\hat{i}j,1990-1980}^{1}) G_{\hat{i}j,2000-1990}^{1}}{P_{i,1980}} \tag{1.13}$$

where $G_{\hat{i}j,1990-1980}^{1}$ and $G_{\hat{i}j,2000-1990}^{1}$ are respectively the growth rates of total number of emigrants during $1980 - 1990$ and $1990 - 2000$ constructed as in equation (1.9). This instrument is uncorrelated with sending countries' specific shocks in 1990 as it uses a prior stock of emigrants in 1980. Also, the growth rates $G^k_{\hat{i}j,2000-1990}$ of migration are computed using a pull factor of migration, such as a total demand for immigrants in destination countries excluding a given migrant sending country, to avoid supply driven shocks.

All instruments have strong explanatory power for all education levels of emigrants except tertiary educated migrants for different country groups, as shown in Table $B.4$. Moreover, the R-squared increases in the first-stage regressions for tertiary educated migrants for non-high income countries and low and lower-middle income countries.

1.6 Estimation Results

This paper studies how the change in the emigration-population ratio affects the growth of GDP per worker. It estimates Equation (1.2) using the OLS and IV approach with two different instruments $IV1$ and $IV2$ to overcome the endogeneity issues. The estimation results are reported in Table $B.5$, where each column is the result of the separate regression and the units of observation are the migrant sending countries. The dependent variable is expressed in terms of the first differences of the logarithm of GDP per worker in the migrant sending countries from 1990 to 2000. The main independent variable is the change in total number of the emigrants relative to the population over the same period of time. "Basic" regressions in Table $B.5$ include only the change in emigration relative to population as an independent variable, while "Extended" regressions add additional control variables to the "Basic" model to test the robustness of the results. The level of the initial dependent variable $(\ln(y_{i,t-1}))$, which is the logarithm of the GDP per worker in 1990, is added to the regressions to consider the catch-up or convergence effects across different income levels of countries. Countries with a low level of GDP per worker grow faster than countries with a high level of GDP per worker, as a growth rate of an economy declines when it approaches its steady state.

Other control variables added to the "Extended" regressions are initial human capital

($s_{i,t-1}$) and average shares of government expenditures and trade in GDP. Countries with higher human capital levels have a higher GDP per worker than others because the marginal product of the highly educated labor is larger than that of the low-skill labor. At the same time, these countries are more likely to grow faster as human capital is a major driving force for productivity improvements.

The average shares of the government expenditures and trade, i.e. imports plus exports, in GDP from 1990 to 2000 also enter the regressions, as a change in these two variables might alter the path of the GDP per worker. The overall impact of higher government spending on GDP per worker depends on several factors including the duration of an increase in public expenditures and government's financing mechanisms (Barro, 1980). While an increase in government spending yields a positive one-to-one change in aggregate demand, it also affects private consumption, thus, changing the overall impact on the GDP. Government spending affects private consumption through three different channels. First, an increase in government spending implies a corresponding increase in taxes, which reduces private consumption unless they are temporary. Second, public spending directly substitutes for private consumption which reduces its impact. Third, increases in taxes on labor earnings lead to a reduced incentive to work and therefore shrink private demand. The main theoretical conclusion is that government spending increases have a major positive impact on GDP if they are temporary. A permanent shift in government spending can affect aggregate demand, but these effects are weak and ambiguous.

The impact of trade policies on GDP and its growth rate depends on model specifications (Rodriguez and Rodrik, 2000). In the static model, trade openness increases GDP due to factor specialization. In standard models with exogenous technological change and diminishing rates of returns there is no impact on GDP growth in the long-run, although there might be growth during a transition to a new policy-driven steady-state. In endogenous growth models trade openness boosts GDP. There might be several gains from trade according to the literature studying monopolistic competition models (Feenstra, 2006). First, an increase in trade brings a price reduction due to increasing returns to scale. As a result of trade liberalization some firms exit the market, while others expand their output and reduce their average costs through economies of scale. Second, more trade expands the product variety available to consumers. Finally, only the most efficient firms survive trade liberalization. These trade effects lead to productivity improvement, thus increasing GDP per capita. However, there might be countries lagging in technological development which would specialize in traditional goods and have lower long-run growth rates due to trade.

These control variables can be potentially correlated with emigration either through GDP per worker or directly. Among direct effects are, for example, a generous public social welfare system or abundant provision of public goods, which might potentially reduce incentives to migrate. These control variables are adopted from Levine et al. (2000) and their list is limited due to the lack of time series for other potential regressors.

All of the above discussed control variables can be potentially a source of endogeneity. First, the estimated Equation 1.2 is expressed in first-differences, and, hence the error term is also expressed in the first-difference and has the following structure: $\Delta\epsilon_{it}^{k} = \epsilon_{it}^{k} - \epsilon_{i,t-1}^{k}$. This error term with a MA(1) process is correlated with the initial dependent variable by construction. Also, the initial human capital might be correlated with the error term $\epsilon_{i,t-1}^{k}$ as there might be factors affecting both GDP per worker and human capital. To correct for this endogeneity bias, 1980's levels of these variables are used as instruments for both the logarithm of GDP per worker and human capital.

Other control variables, such as, shares of government expenditures and trade in GDP, can also be a potential source of bias caused by reverse causality and endogeneity. For instance, higher government expenditures increase GDP per worker, which might cause higher government expenditures as more economic activities leads to higher taxes and, hence, more resources for the government to spend. Also, there is a correlation between the trade shares and the error term as trade is a function of GDP, so there might be other factors affecting both trade and GDP per worker. To address these problems the paper instruments these variables by their averages over 1980–1990 in all IV estimates.

1.6.1 Results for All Countries

Emigration positively affects GDP per worker in most econometric specifications, but the results are weak. A change in total emigration relative to population increases GDP per worker by 0.8 in OLS estimation, as shown for the ”Basic” model in Table $B.5$. This coefficient is significant at the 10 percent level and retains its significance and sign with the inclusion of control variables. $IV1$ estimations also produce significant and positive results but with a higher magnitude of about 1.4 at the 10 percent significance level for both ”Basic” and ”Extended” specifications. $IV2$ estimations yield insignificant results. Among control variables in the ”Extended” specifications, the initial log of GDP per worker and human capital are significant in all estimations at either the 5 or the 1 percent level. The only exception is the initial GDP per worker in the OLS specification. Both of these variables have the expected signs: negative for GDP per worker and positive for human capital. The coefficients are insignificant for other control variables across various estimates, except for the average government expenditures in the OLS estimates, which has a significant coefficient of -0.01. Thus, GDP per worker increases in response to total emigration changes as a percentage of population both in the OLS and $IV1$ approaches, but these results are are weak with a significance level at 10 percent at most.

1.6.2 Breaking Down by Education Categories

Emigrants with different education levels can have diverse effects on the economies of migrant sending countries. For example, emigration of both low-skill and high-skill workers shrinks the labor force, but might have both a positive and negative impact on GDP per worker dependent on the skill level of emigrants. Emigration of low-skill labor would not necessarily reduce GDP per worker in the sending country. If the composition of the labor force shifts towards having less low-skill and more high-skill labor, then GDP per worker would increase due to a higher marginal product of more educated as opposed to less educated workers. At the same time, emigration of high skill labor might result in the reduction of GDP per worker if one considers only the direct impact on the labor force. Emigration of more educated workers can also have a positive impact on GDP per worker by increasing the average level of human capital in the country. Emigration rates are higher for secondary and tertiary educated people (0.051) and for tertiary educated people (0.15) than for the total number of emigrants (0.013), as shown in Table $B.3$. Higher emigration rates for more educated people mean a higher likelihood to emigrate for the potential pool of migrants. This leads to higher investments in education and increases the average human capital in the country as not everyone has a chance to emigrate. This channel of emigration impact on GDP per worker is specific only to high-skill labor and has been widely studied in the literature (Beine et al, 2008; Docquier et al., 2008).

Education is also an important factor in terms of migrants' destinations. More educated migrants move to more advanced economies which have been increasingly implementing policies to attract high-skill labor. Having diasporas in different countries might help promote bilateral relationships between migrant sending and receiving countries and contribute to larger trade flows, transfer of know-how, etc. These effects might become stronger with an increase in education levels of emigrants and their relocation to frontier economies with sizable economies and a high level of technological advancement.

Table $B.6$ reports estimation results for the impact of different education categories of emigrants on GDP per worker which vary across different econometric specifications. Only a change in the total emigration relative to population has a positive impact on the growth of GDP per worker in both OLS and $IV1$ estimations at the 10 percent significance level. In contrast, $IV2$ estimations yield insignificant coefficients. Both secondary and tertiary educated migrants, and tertiary educated migrants have no significant impact on GDP per worker with an exception of tertiary educated migrants in the "Extended" model in OLS (0.04) and $IV1$ (0.08) estimations at the 10 percent significance level. Thus, the estimation results do not improve when broken down by education categories.

1.6.3 Breaking Down by Country Groups and Education Categories

Labor mobility can have different economic implications for various income groups of countries. For instance, emigration of high-skill labor in poor developing countries might be more detrimental for their economies than in high-income countries. The skilled labor in many developing countries is a rare production input with a much higher marginal product compared to other labor inputs. Hence, the emigration of skilled workers in these circumstances might shrink GDP per worker more than in the advanced economies with smaller gaps in marginal products across different labor inputs. Also, migrants are a selected group of people who demonstrate more risk-taking behavior than does the average individual given their decisions to reallocate to foreign destinations with a lot of uncertainty. The outflow of this selected group of population might reduce the critical mass of people who are key to technological progress and innovation. In poor countries with low productivity levels this can have larger implications than in developed economies, which are at the frontier of technological advancement.

Labor migration can also have varying positive effects across different income groups of countries. Large migrant diasporas are often associated with better bilateral relationships between migrant sending and migrant hosting countries. These relationships may promote bilateral trade and foreign direct investment which can raise GDP per worker by creating a larger demand and inducing technology transfer between countries (Gould, 1994). For poor countries with low level productivity these effects might be much larger than for more advanced economies. Also, diasporas might generate more capital flows to countries of origin that are more in need than others. In addition, closer ties with high-income countries and exposure to a better way of doing business might bring improvement in institutions and, therefore, an increase in GDP. Again these effects might be smaller for countries with higher levels of income than others. Finally, migrants from developing countries might remit more money to their home countries given more severe financial constraints of their households, relatives, and friends than migrants from higher income countries. At the same time, the marginal effects of remittances might be larger in poor economic conditions than in rich countries. The list of factors driving potential differences between high-income and low-income countries is not exhaustive here and the main purpose is to provide some intuition for possible differences in impact of emigration across different income groups.

Table $B.7$ extends the present framework to explore emigration effects for different income groups of countries. It reports regression results of emigration impact on GDP per worker for different education categories and for five groups of countries: (1) all countries, (2) non-high income countries; (3) low and lower-middle income countries; (4) high income countries; and (5) upper-middle income countries. The focus of this discussion are non-high income countries and low and lower-middle income countries. The results for other country groups are insignificant with an exception of the negative impact of the total emigration in high income countries. The rationale of focusing on these two country groups, first, is driven by the existence of sharp differences in economic incentives to migrate between high income countries, and high-income and non-high income countries. There are either no economic incentives for migration between high-income countries or they are relatively weak, and migration is primarily driven by other factors such as preferences of individuals. At the same time, within the non-high income country group, low and lower-middle income countries are the poorest economies where migration effects can be much larger given the scarcity of human capital and high financial constraints.

The magnitude and significance of the impact of a change in the total emigration relative to population for these two country groups tend to increase. In particular, both

OLS and $IV1$ estimates show a positive impact of total emigration on GDP per worker growth with coefficients in the range of 1.14–1.55 for non-high income countries. All coefficients are larger in magnitude compared to estimates for all countries with an exception of the "Basic" model in $IV1$ with a coefficient of 1.18 as opposed to 1.43 for all countries. This is the only estimate which retains its significance level at 10 percent, while all other coefficients become significant at the 5 percent level. The only significant $IV2$ estimate is in the "Extended" model of non-high income countries with a coefficient of 1.1 at a significance level of 10 percent. The estimation results become stronger both in terms of the magnitude of coefficients and their significance for low and lower-middle income country groups. All coefficients are significant, within the range of $1.7 - 2.29$, and robust to the inclusion of different control variables. This indicates that economic implications of emigration are much stronger for less developed countries. Considering that the median changes in the total number of emigrants relative to population in non-high income countries and low and lower-middle income countries are respectively 0.007 and 0.004, an increase in emigration relative to population from zero to its median value would raise the growth rate of GDP per worker by nearly one percentage point in both country groups. While these effects might seem high, they reflect decade long changes in GDP per worker, so that the change on an annual basis is approximately one tenth of one percentage point. Although the coefficients are larger for low and lower-middle income countries, the growth is similar in two country groups given the high median change in emigration relative to population in non-high income countries as opposed to low and lower-middle income countries. Lower emigration rates in low and lower-middle income countries compared to others can be a reflection of financial constraints faced especially by low-skill labor which impede costly mobility across countries. The impact of emigration also becomes significant and positive for $IV2$ estimations in the "Extended" model for non-high income countries and in both "Basic" and "Extended" specifications for low and lower-middle income countries with coefficients similar to $IV1$.

The impact of emigrants with secondary and tertiary education varies across different specifications. GDP per worker grows in response to changes in the number of secondary and tertiary educated emigrants relative to their respective population in the ”Extended” $IV1$ model for non-high income countries, and in the ”Basic” model for both OLS and $IV1$ specifications and in the ”Extended” $IV1$ estimations for low and lower-middle income countries. The estimates produce more significant results for low and lower-middle income countries as opposed to non-high income countries, while emigration has no impact when all countries are included in the sample. As in the case of total emigration, the significant coefficients have higher positive values for low and lower-middle income countries (0.37–0.54) than for non-high income countries (0.3). The only robust estimate is IV1 for low and lower-middle income countries which remains both positive and significant in both ”Basic” (0.46) and ”Extended” (0.54) specifications. If this estimate is used, then the median change in the number of secondary and tertiary educated emigrants relative to population would bring nearly 2 percent growth in GDP per worker. The magnitude of the coefficients for secondary and tertiary educated emigrants is lower than for the total number of emigrants, possibly reflecting the significance of education independent channels. For instance, transmission channels of emigration such as diaspora-induced trade or foreign direct investment are not conditional on education.

Finally, both OLS and $IV1$ estimations produce significant and positive coefficients for the change in the number of emigrants with tertiary education as a percentage of population on GDP per worker for non-high income countries and low and lower-middle income countries. These coefficients range from 0.04 to 0.1 and are much smaller in magnitude than those for secondary and tertiary educated emigrants and for the total number of emigrants. Using the median values for the change in the number of emigrants with tertiary education relative to population and the coefficients of $IV1$ estimates, the response of the growth in the GDP per worker will be 1 and 2 percent respectively for non-high income and low and lower-middle income countries. The effects for low and lower-middle income countries are larger than for non-high income countries due to both larger coefficients and median change in the number of emigrants relative to population.

One can test for a difference in the coefficients of the emigration change relative to population between various country groups by including an interaction term in the regressions. The paper estimates all "Basic" IV regressions for the entire sample of countries by including a dummy variable for low and lower-middle income countries and an interaction term of this dummy variable and a change in emigration relative to population. It focuses on the low and lower-middle income countries as they have the most robust significant estimates across different specifications. The interaction term is endogenous given the endogeneity of the emigration variables and the natural instrument in this case would be an interaction of the dummy variable and the respective instrument for the emigration change relative to population (Wooldridge, 2002). The paper estimates these regressions for all education groups. The results indicate that the only significant coefficient for the interaction term among different education groups is for $IV2$ estimates for total emigration. Thus, only in this case the coefficients between lower-middle income countries and others are significantly different.

1.6.4 Decomposing Effects

Decomposing the GDP per worker into capital per worker, human capital, and TFP helps understand the main channels of GDP growth driven by emigration. Emigrants might potentially affect all components of GDP per worker. Overall, developing countries have low levels of capital due to financial constraints and a low level of savings. Remittances sent home by emigrants might relax these financial constraints and lead to physical capital accumulation in the capital scarce environment. Table $B.2$ shows that despite a low level of capital per worker in non-high income and low and lower-middle income countries, the median growth rates of the capital-worker ratio is much lower in those countries than in all countries. Thus, in the capital constrained conditions migrants' remittances can be essential for physical capital accumulation.

Another channel of emigration impact on GDP per worker can work through human capital. There is a huge empirical literature discussing the impact of emigration on the human capital of migrant sending countries. Due to immigration policies in major migrant destination countries, such as the U.S., Canada, and Australia, aimed to attract high-skill labor, the likelihood to migrate is higher for educated individuals than for people without education. As emigrants receive higher compensation for their services, this increases the marginal returns to education and, hence, investments in education. As not everyone can emigrate, the overall human capital in the country improves.

Finally, emigration may affect the GDP per worker through TFP. Productivity improvements can be driven by both direct and indirect channels of technology transfer between migrant sending and receiving countries. For example, direct transfers can take place through the foreign direct investments promoted by emigrants. Indirect channels of productivity improvements can be trade facilitated by emigrants, which induces more competition or exposure to better institutions for migrant sending countries.

Estimating the impact of emigration on the capital-worker ratio for non-high income and low and lower-middle income countries produces qualitatively different results as opposed to insignificant estimates for all countries. Table $B.8$ demonstrates that the $IV1$ estimates in the "Extended" model are significant and positive for all education categories in low and lower-middle income countries with a coefficient of 0.02 for tertiary educated emigrants; 0.12 for emigrants with secondary and tertiary educated emigrants, and 0.55 for the total number of emigrants. Considering that the median changes in emigration relative to population are 0.2, 0.044, and 0.004 respectively for tertiary educated, secondary and tertiary educated, and all emigrants, the growth rates of the capital-worker ratio would be 0.4, 0.5, and 0.2 percent. $IV1$ estimates produce similar results for non-high income countries with significant coefficients for the total number of emigrants and emigrants with secondary and tertiary education. There is no impact of tertiary educated emigrants on the capital accumulation in non-high income countries.

Table $B.9$ shows no significant impact of the change in emigration on human capital across different education categories and country groups. The only significant positive result at the 10 percent significance level is for total emigration in the "Extended" $IV1$ model. Thus, this is contrary to the findings of the related literature on the positive impact of emigration on the human capital of migrant sending countries.

The main component driving the growth in GDP per worker is TFP as shown in Table $B.10$. Similar to GDP per worker, all estimation results for low and lower-middle income countries are significant and positive in the "Basic" model both for OLS and $IV1$ and are in the range of 0.061–1.78. These results are robust to the inclusion of control variables in all specifications. Considering the median changes in emigration relative to population for three education categories, TFP growth would account for 1 percent for total number of emigrants, 1.9 percent for emigrants with secondary and tertiary education, and 1.7 percent for tertiary educated emigrants. These effects are stronger for more educated emigrants due to high median change in emigration relative to population despite smaller coefficients. Among other estimates the $IV1$ estimations produce significant positive results for all emigrants in the "Extended" model in non-high income countries as in the case of all countries.

Thus, TFP is a major source for improvement in GDP per worker in low and lower-middle income countries, but the education level is not the underlining reason for the increase in productivity. A higher growth in GDP per worker in response to the median change in emigration relative to population for tertiary educated emigrants than for other groups is driven by greater changes in their emigration rates rather than by larger coefficients. These changes in TFP might be a result of trade or other cross-country partnerships facilitated by established diasporas abroad which lead to a transfer of knowledge and technology.

1.7 Conclusions

This paper studies the impact of the change in emigration relative to population on the logarithm of the GDP per worker of migrant-sending countries, expressed as growth rates. It uses 1990 and 2000 emigration data from 195 source countries to 30 OECD destination countries and adopts a migrant sending country as the unit of observation. This paper contributes to the existing literature by extending the current limited research on the impact of the brain drain on the growth of migrant sending countries. It discusses the impact of emigrants with three different education levels: all emigrants, emigrants with secondary and tertiary education, and emigrants with only tertiary education, on GDP per worker in different income groups of countries. Also, it suggests a new set of instruments which provide exogenous variations in emigration and have strong explanatory power as opposed to the previous ones used in the literature.

The paper uses new instruments to address the endogeneity and simultaneity bias in OLS estimations. These instruments are based on the migration pull factors such as demand for migrants in destination countries and migrants' networks. First, the paper computes the growth rates of immigration in 30 OECD countries to determine the demand for immigrants by destination countries. As the changes in immigration are primarily driven by immigration policies and labor demand in migrant receiving countries, these growth rates can be treated as exogenous to the conditions in countries of origin. Next, the growth rate of immigration in a given destination country is applied to the number of immigrants from different countries of origin in 1980. As migrants' networks play an important role in location choices of migrants, it is assumed that an increase in the number of immigrants in a given destination country would be proportional to the sizes of their diasporas. The year 1980 is used instead of 1990 to avoid endogeneity problems caused by a correlation between 1990 variables and the error term, which is a MA(1) process by construction. Finally, the instrument is computed as a difference of the constructed number of migrants aggregated across destination countries and the actual number of emigrants in 1990.

Estimation results indicate that an increase in the total emigration relative to the population raises GDP per worker with a coefficient of 2 in low and lower-middle income countries. These labor productivity improvements are primarily driven by increases in TFP possibly through trade, FDI, and other cross-country partnerships facilitated by emigrants' diasporas leading to a transfer of knowledge and technology. The results are robust to the inclusion of various control variables in the regressions.

Bibliography

[1] Acosta, P., C. Calderon, P. Fajnzylber and H. Lopez (2008): What is the Impact of International Remittances on Poverty and Inequality in Latin America?, World Development, 36, 1: 89-114.

[2] Adams, R. H. Jr. (1998): Investment, and Rural Asset Accumulation in Pakistan, Economic Development and Cultural Change, 47, 1: 155-173.

[3] Aydemir, A., and G. J. Borjas (2007): A comparative analysis of the labor market impact of international migration: Canada, Mexico, and the United States, Journal of the European Economic Association, 5, 4: 663-08.

[4] Barajas, A., R. Chami, C. Fullenkamp, M. Gapen and P. Montiel (2009): Do Workers' Remittances Promote Economic Growth?, IMF Working Paper, WP/09/153.

[5] Barro, R.J. (1980): Output Effects of Government Purchases, National Bureau of Economic Research, Working Paper 432.

[6] Barro, R.J. and J.W. Lee (2000): International Data on Educational Attainment: Updates and Implications, CID Working Papers 42, Center for International Development (Harvard University).

[7] Beine, M., F. Docquier and C. Oden-Defoort (2011): A Panel Data Analysis of The Brain Gain, World Development.

[8] Beine, M., F. Docquier and C. Ozden (2011): Diasporas, Journal of Development Economics, 95, 1: 30-41.

[9] Beine, M., F. Docquier and H. Rapoport (2001): Brain drain and economic growth: theory and evidence, Journal of Development Economics, 64, 1: 275-89.

[10] Beine, M., F. Docquier and H. Rapoport (2007): Measuring international skilled migration: new estimates controlling for age of entry, World Bank Economic Review, 21, 2: 249-54.

[11] Beine, M., F. Docquier and H. Rapoport (2008): Brain drain and human capital formation in developing countries: winners and losers, Economic Journal, 118: 631-652.

[12] Card, D. (2001): Immigrant Inflows, Native Outflows, and the Local Market Impacts of Higher Immigration, Journal of Labor Economics 19: 22-64.

[13] Catrinescu, N., M. Leon-Ledesma, M. Piracha and B. Quillin (2009): Remittances, Institutions, and Economic Growth, World Development 37, 1: 81-92.

[14] Caselli, F. (2005): Accounting for Cross-Country Income Differences, Handbook of Economic Growth, in: Philippe Aghion and Steven Durlauf (ed.), Handbook of Economic Growth, edition 1, volume 1, chapter 9: 679-741. Elsevier.

[15] Chami, R., D. Hakura and P. Montiel (2009): Remittances: An Automatic Output Stabilizer?, IMF Working Paper, WP/09/91.

[16] Cohen, D. and M. Soto (2001): Growth and human capital: good data, good results, CEPR Discussion Paper, n. 3025.

[17] Cox, A. and M. Ureta (2003): International migration, remittances and schooling: evidence from El Salvador, Journal of Development Economics, 72, 2: 429-461.

[18] De la Fuente, A. and R. Domenech (2002): Human capital in growth regressions: how much difference does data quality make? An update and further results, CEPR Discussion Paper, n. 3587.

[19] Docquier, F., O. Faye and P. Pestieau (2008): Is migration a good substitute for subsidies, Journal of Development Economics, 86, 2: 263-76.

[20] Docquier, F., O. Lohest and A. Marfouk (2007): Brain drain in developing countries, World Bank Economic Review, 21, 2: 193-218.

[21] Docquier, F., B. L. Lowell and A. Marfouk (2009): A gendered assessment of the brain drain, Population and Development Review, 35, 2: 297-321.

[22] Docquier, F. and A. Marfouk (2006): International migration by educational attainment (1990-2000), in C. Ozden and M. Schiff (eds): International Migration, Remittances and Development, Palgrave Macmillan: New York.

[23] Easterly, W. and Y. Nyarko (2009): Is the Brain Drain Good for Africa? Chapter 11 in J. Bhagwati and G. Hanson (eds): Skilled immigration: problems, prospects and policies, Oxford University Press: 316-60.

[24] Feenstra, R. C. (2006): New Evidence on the Gains from Trade, Review of World Economics, 142: 617-641.

[25] Gould, D. M. (1994): Immigration links to the home country: Empirical implications for U.S. bilateral trade flows, Review of Economics and Statistics, 76: 302-316.

[26] Grogger, J. and G. Hanson (2011): Income maximization and the selection and sorting of international migrants, Journal of Development Economics, 95, 1: 42-57.

[27] Grubel, H. and A. Scott (1966): The international flow of human capital, American Economic Review, 56: 268-74.

[28] Hall, R.E., and Jones, C.I. (1999): Why do some countries produce so much more output per worker than others?,The Quarterly Journal of Economics 114 (1), 83–116.

[29] Hanson, G. H. (2007): International Migration and the Developing World, Mimeo, UCSD.

[30] Head, K., J. Ries and D. Swenson (1998): Immigration and trade creation: Econometric evidence from Canada, Canadian Journal of Economics, 31, 1: 47-62.

[31] Hildebrandt, N. and D. McKenzie (2005): The effects of migration on child health in Mexico, World Bank Policy Research Working Paper, 3573.

[32] Levine, R., N. Loayza and T. Beck (2000): Financial intermediation and growth: Causality and causes, Journal of Monetary Economics, 46: 31-77.

[33] Mayda, A.M. (2010): International migration: a panel data analysis of the determinants of bilateral flows, Journal of Population Economics, 23, 4: 1249-74.

[34] Mishra, P. (2007): Emigration and wages in source countries: Evidence from Mexico, Journal of Development Economics, 82, 1: 180-199.

[35] Nielsen, M. E., and P. Olinto (2007): Do Conditional Cash Transfers Crowd Out Private Transfers? in P. Fajnzylber and H. Lopez (eds.): Remittances and Development: Lessons from Latin America, The World Bank, Washington, DC: 253-298.

[36] Rodriguez, F. and D. Rodrik (2001): Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence, NBER Macroeconomics Annual 2000, Volume 15.

[37] Saxenian, A. (1999): Silicon Valley's new immigrant entrepreneurs, Public Policy Institute of California.

[38] Teruel, G. and B. Davis (2000): Final report: An evaluation of the impact of PROGRESA cash payments on private inter-household transfers, International Food Policy Research Institute.

[39] Schumpeter, J. A. (1934): The Theory of Economic Development, New York, Oxford University Press.

[40] Woodruff, C., and R. Zenteno (2007): Migration networks and microenterprises in Mexico, Journal of Development Economics, 82, 2: 509-528.

[41] Wooldridge, J.M. (2002): Econometric analysis of cross-section and panel data, The MIT Press, Cambridge.

[42] World Bank Group and Ç. Özden, C. Parsons, M. Schiff and T. L. Walmsley (2011): Global Bilateral Migration Database, Where on Earth is Everybody? The Evolution of Global Bilateral Migration, 1960-2000, World Bank Economic Review 25(1):12-56.

[43] Yang, D. (2008): International migration, remittances, and household investment: Evidence from Philippine migrants' exchange rate shocks, Economic Journal, 118, 528: 591-630.

Appendix A

Data Appendix

This section provides sources and description of the variables:

1. Emigration variables are from dataset by Docquier, Lowell, and Marfouk (2008) which provides the number of migrants from 195 migrant-sending countries to 30 main destination OECD countries in 1990 and 2000. The main migration variable is $E_{it}^{k}/P_{i,t-1}^{k}$ where E_{it}^{k} is number of emigrants from country i in period t for education group k and $P_{i,t-1}^{k}$ is a population in country i in period $t-1$ which measures a change in emigration relative to population is constructed for three different education groups: all emigrants, emigrants with secondary and tertiary education, and emigrants with tertiary education. The paper uses labor force data from the same dataset instead of population for secondary and tertiary educated migrants.

2. Employment-population ratio, GDP per worker, investments, number of workers, shares of government expenditures and trade in GDP, and population data are obtained from the Penn World Tables (PWT) by Heston, Summers and Bettina (PWT 7.0). The number of workers in each country i and year t is computed as ($rgdpch_{it} * pop_{it}/rgdpwok_{it}$), where $rgdpch_{it}$ is a PPP converted GDP per capita (Chain Series) at 2005 constant prices, pop_{it} is a population, and $rgdpwok_{it}$ is a PPP Converted GDP Chain per worker at 2005 constant prices.

3. The average human capital h_{it} is constructed using average years of schooling in the population over 25 years old from the Barro-Lee dataset (Barro and Lee, 2000). As in Docquier and Marfouk (2006), human capital indicators are replaced with those from De La Fuente and Domenech (2002) for OECD countries. For countries where Barro and Lee measures are missing, the proportion of educated individuals is predicted using the Cohen and Soto (2007) measures.

Appendix B

Tables and Figures

Table B.1: List of countries by Country Groups

Non-high Income	**Low and Lower-Middle Income**
Afghanistan	Afghanistan
Albania	Angola
Algeria	gladesh
Angola	Belize
Argentina	Benin
Bangladesh	Bhutan
Belize	Bolivia
Benin	Burkina Faso
Bhutan	Burundi
Bolivia	Cambodia
Bosnia and Herzegovina	Cameroon
Botswana	Cape Verde
Brazil	Central African Republic
Bulgaria	Chad
Burkina Faso	Comoros
Burundi	Congo, DRC
Cambodia	Congo, Republic
Cameroon	Cote d'Ivoire
Cape Verde	Djibouti
Central African Republic	Egypt

Chad	El Salvador
Chile	Ethiopia
China	Fiji
Colombia	The Gambia
Comoros	Ghana
Congo, DRC	Guatemala
Congo, Republic	Guinea
Costa Rica	Guinea-Bissau
Cote d'Ivoire	Guyana
Cuba	Haiti
Djibouti	Honduras
Dominican Republic	India
Ecuador	Indonesia
Egypt	Iraq
El Salvador	Kenya
Ethiopia	Laos
Fiji	Lesotho
Gabon	Liberia
The Gambia	Madagascar
Ghana	Malawi
Grenada	Mali
Guatemala	Mauritania
Guinea	Mongolia
Guinea-Bissau	Morocco
Guyana	Mozambique
Haiti	Nepal
Honduras	Nicaragua
India	Niger
Indonesia	Nigeria
Iran	Pakistan
Iraq	Papua New Guinea
Jamaica	Paraguay
Jordan	Philippines
Kenya	Rwanda

Laos	Samoa
Lebanon	Sao Tome and Principe
Lesotho	Senegal
Liberia	Sierra Leone
Libya	Solomon Islands
Macedonia	Somalia
Madagascar	Sri Lanka
Malawi	Sudan
Malaysia	Swaziland
Maldives	Syria
Mali	Tanzania
Mauritania	Togo
Mauritius	Tonga
Mexico Mongolia	Uganda Uzbekistan
Morocco	Vanuatu
Mozambique	Vietnam
Namibia	Yemen
Nepal	Zambia
Nicaragua	Zimbabwe
Niger	
Nigeria	
Pakistan	
Panama	
Papua New Guinea	
Paraguay	
Peru	
Philippines	
Romania	
Russia	
Rwanda	
Saint Lucia	
St. Vincent and the Grenadines	
Samoa	

Sao Tome and Principe	
Senegal	
Seychelles	
Sierra Leone	
Solomon Islands	
Somalia	
South Africa	
Sri Lanka	
Sudan	
Suriname	
Swaziland	
Syria	
Tanzania	
Thailand	
Togo	
Tonga	
Tunisia	
Turkey	
Uganda	
Uruguay	
Uzbekistan	
Vanuatu	
Venezuela	
Vietnam	
Yemen	
Zambia	
Zimbabwe	

Table B.2: Summary Statistics: Median and Mean Growth Rates of Dependent Variables by Country Groups

	All Income		**Non-High Income**		**Low and Lower-Middle Income**	
	Median	**Mean**	**Median**	**Mean**	**Median**	**Mean**
GDP per Worker Growth	0.11	0.16	0.04	0.07	-0.02	0.02
Capital per Worker Growth	0.04	0.11	0.03	0.06	0.02	0.03
Human Capital Growth	0.07	0.07	0.08	0.09	0.08	0.09
TFP Growth	-0.03	-0.02	-0.08	-0.07	-0.12	-0.11
Number of Observations	162	162	115	115	74	74

Table presents the growth rates of the GDP per worker and its components by country income groups. The presented data are median and mean growth rates for each group over $1990 - 2000$.

Table B.3: Summary Statistics: Median and Mean of Emigration Variables by Country Groups

	All Income		**Non-High Income**		**Low and Lower-Middle Income**	
	Median	**Mean**	**Median**	**Mean**	**Median**	**Mean**
$\Delta E^1_{it}/P^1_{i,t-1}$	0.006	0.025	0.007	0.03	0.004	0.026
$E^1_{it}/P^1_{i,t}$	0.025	0.073	0.013	0.071	0.007	0.053
$\Delta E^2_{it}/P^2_{i,t-1}$	0.033	0.107	0.042	0.135	0.044	0.124
$E^2_{it}/P^2_{i,t}$	0.054	0.17	0.051	0.196	0.046	0.156
$\Delta E^3_{it}/P^3_{i,t-1}$	0.09	0.435	0.13	0.555	0.2	0.585
$E^3_{it}/P^3_{i,t}$	0.138	0.541	0.15	0.646	0.186	0.6
Number of Observations	162	162	115	115	74	74

Table presents the change in emigration relative to population - $\Delta E^k_{it}/P^k_{i,t-1}$, and emigration-population ratio - $E^k_{it}/P^k_{i,t}$ by country income groups. The data for the change in emigration relative to population are median and mean changes for $1990-2000$ and data for the emigration-population ratio are medians and means for 2000.

Table B.4: First-Stage Regression Results

Dependent Variable	**Coeff.**	**t-stat**	**R-squared**	**Obs.**
All Countries				
All Emigrants: IV1	0.44	7.36	0.26	162
All Emigrants: IV2	0.58	8.21	0.3	162
Secondary and Tertiary Emigrants: IV1	1.09	5.59	0.16	162
Tertiary Emigrants: IV1	3.15	3.66	0.08	162
Non-High Income Countries				
All Emigrants: IV1	0.52	7.64	0.34	115
All Emigrants: IV2	0.76	9.21	0.43	115
Secondary and Tertiary Emigrants: IV1	1.51	5.93	0.24	115
Tertiary Emigrants: IV1	4.74	3.97	0.12	115
Low and Lower-Middle Income Countries				
All Emigrants: IV1	0.71	9.21	0.54	74
All Emigrants: IV2	1.35	15.9	0.78	74
Secondary and Tertiary Emigrants: IV1	2.77	12.74	0.69	74
Tertiary Emigrants: IV1	14.9	6.63	0.38	74

Table presents first-stage $IV1$ regression results for the change in emigration relative to population for three different education groups of emigrants and $IV2$ regression results for the total number of emigrants for different country groups. The independent variable in $IV1$ is Z_{1i}^{k}, while the independent variable in $IV2$ is Z_{2i}^{1}. $IV2$ is applied only for the total number of emigrants as there are no data on education of emigrants in 1980 to construct the growth rates for immigration in destination countries for the period of $1980 - 1990$.

Table B.5: Second-Stage Regression Results: GDP per Worker, All Countries

Regressors	OLS		IV1		IV2	
	Basic	**Extended**	**Basic**	**Extended**	**Basic**	**Extended**
$\frac{\Delta E^1_{it}}{P^1_{i,t-1}}$	0.8*(0.36)	1.02*(0.4)	1.43*(0.71)	1.42*(0.62)	1.14(0.76)	0.99(0.65)
$\ln(y_{i,t-1})$		-0.05(0.05)		-0.1**(0.04)		-0.1**(0.04)
$s_{i,t-1}$		0.51**(0.18)		0.82***(0.2)		0.83***(0.2)
Average government expenditures		-0.01*(0.01)		-0.01(0.01)		-0.01(0.01)
Average trade share in GDP		0(0)		0(0)		0(0)
Number of observations	162	162	162	147	162	147
R-Squared	0.01	0.1				

Each column is the result of a separate regression where the units of observations are migrant sending countries with a change in respective variables over the period of $1990-2000$. The Table presents regression results for the impact of a change in the total number of emigrants relative to population - $\frac{\Delta E^1_{it}}{P_{i,t-1}}$, on the GDP per worker growth measured as a derivative of its logarithm. The methods of estimation are Ordinary Least Squares (OLS) and Instrumental Variable (IV) approach with two different instruments $IV1$ and $IV2$. "Basic" regressions include only the change in emigration relative to population as an independent variable while the "Extended" regressions add to 'Basic' additional control variables: initial dependent variable $(\ln(y_{i,t-1}))$ and human capital $(s_{i,t-1})$, average of government expenditures and trade shares in GDP. All control variables in the IV regressions are instrumented by their lags to address the endogeneity issues. The numbers in parentheses are heteroskedasticity robust standard errors of the coefficients and (*) indicates significance level at 10, (**) at 5, and (***) at 1 percent.

Table B.6: Breaking down by Education Categories: GDP per worker, All Countries

Regressors	OLS		IV1		IV2	
	Basic	**Extended**	**Basic**	**Extended**	**Basic**	**Extended**
$\frac{\Delta E^1_{it}}{P^1_{i,t-1}}$	0.8*(0.36)	1.02*(0.4)	1.43*(0.71)	1.42*(0.62)	1.14(0.76)	0.99(0.65)
$\frac{\Delta E^2_{it}}{P^2_{i,t-1}}$	0.09(0.09)	0.15(0.11)	0.36(0.22)	0.29(0.16)		
$\frac{\Delta E^3_{it}}{P^3_{i,t-1}}$	0.03(0.02)	0.04*(0.02)	0.11(0.06)	0.08*(0.04)		

Each cell is the result of a separate regression where the units of observations are migrant sending countries with a change in respective variables over the period of $1990 - 2000$. The Table presents regression results for the impact of a change in the total number of emigrants relative to population - $\frac{\Delta E^k_{it}}{P^k_{i,t-1}}$, on the GDP per worker growth measured as a derivative of its logarithm. k superscript denotes an educational category which equals 1 for total number of migrants, 2 for migrants with secondary and tertiary education, and 3 for migrants with tertiary education. The methods of estimation are Ordinary Least Squares (OLS) and Instrumental Variable (IV) approach with two different instruments $IV1$ and $IV2$. $IV2$ is constructed only for total number of migrants. ”Basic” regressions include only the change in emigration relative to population as an independent variable while the ”Extended” regressions add to ’Basic’ additional control variables: initial dependent variable and human capital, average of government expenditures and trade shares in GDP. All control variables in the IV regressions are instrumented by their lags to address the endogeneity issues. The numbers in parentheses are heteroskedasticity robust standard errors of the coefficients and (*) indicates significance level at 10, (**) at 5, and (***) at 1 percent.

Table B.7: Breaking down by Country Groups and Education Categories: GDP per worker

	All		**Non-High Income**		**Low, Lower-Mid. Income**		**High Income**		**Upper-Middle Income**	
	Basic	**Extended**	**Basic**	**Extended**	**Basic**	**Extended**	**Basic**	**Extended**	**Basic**	**Extended**
$\frac{\Delta E^1_{it}}{P^1_{i,t-1}}$: OLS	0.8*(0.36)	1.02*(0.4)	1.14**(0.39)	1.28**(0.44)	1.67**(0.52)	1.78*(0.72)	0.46(1.27)	-2.89**(1.04)	0.05(0.45)	0.22(0.54)
$\frac{\Delta E^1_{it}}{P^1_{i,t-1}}$: IV1	1.43*(0.71)	1.42*(0.62)	1.18*(0.5)	1.55**(0.46)	1.79***(0.49)	2.29***(0.52)	-4.94(15.63)	-22.02(44.96)	-1.17(1.36)	0.47(0.62)
$\frac{\Delta E^1_{it}}{P^1_{i,t-1}}$: IV2	1.14(0.76)	0.99(0.64)	0.84(0.53)	1.1*(0.55)	1.7***(0.46)	2.12***(0.6)	-3.22(10.94)	-15.57(20.85)	-1.61(1.3)	-0.21(0.91)
$\frac{\Delta E^2_{it}}{P^2_{i,t-1}}$: OLS	0.09(0.09)	0.15(0.11)	0.13(0.1)	0.16(0.1)	0.37*(0.16)	0.33(0.19)	1.91(1.72)	-0.06(0.58)	-0.05(0.12)	0.02(0.11)
$\frac{\Delta E^2_{it}}{P^2_{i,t-1}}$: IV1	0.36(0.22)	0.29(0.16)	0.25(0.15)	0.3*(0.14)	0.46***(0.12)	0.54***(0.14)	-0.27(1.54)	-1.67(1.41)	-0.4(0.38)	-0.05(0.18)
$\frac{\Delta E^3_{it}}{P^3_{i,t-1}}$: OLS	0.03(0.02)	0.04*(0.02)	0.04*(0.02)	0.05*(0.02)	0.05*(0.02)	0.05*(0.02)	0.1(0.06)	-0.12(0.07)	0.01(0.02)	0.03(0.03)
$\frac{\Delta E^3_{it}}{P^3_{i,t-1}}$: IV1	0.11(0.06)	0.08*(0.04)	0.07*(0.03)	0.07**(0.02)	0.1***(0.02)	0.1**(0.03)	-0.05(0.35)	-0.39(0.27)	-0.11(0.16)	0(0.04)

Each cell is the result of a separate regression where the units of observations are migrant sending countries in 2000. Table presents regression results for the impact of the growth in three different groups of emigrants relative to population: all emigrants (All), emigrants with secondary and tertiary education (Secondary and Tertiary), and emigrants with tertiary education (Tertiary) on GDP growth. The methods of estimation are Ordinary Least Squares (OLS) and Instrumental Variable (IV) approach with different instruments constructed for each group of emigrants.'Basic' regressions include only the growth in emigration relative to population as an independent variable while the 'Extended' regressions add to 'Basic' additional control variables: initial dependent variable and human capital, average of government expenditures and trade share in GDP. All control variables in IV regressions are instrumented by their lags to ensure their exogeneity. The numbers in parentheses are heteroskedasticity robust standard errors of the coefficients and (*) indicates significance level at 10, (**) at 5, and (***) at 1 percent.

Table B.8: Decomposing Effects: Capital per Worker

	All Countries		Non-High Income		Low and Lower-Middle Income	
	Basic	**Extended**	**Basic**	**Extended**	**Basic**	**Extended**
$\frac{\Delta E^1_{it}}{P^1_{i,t-1}}$: OLS	0.11(0.11)	0.18(0.14)	0.22(0.11)	0.27(0.14)	0.08(0.11)	0.18(0.17)
$\frac{\Delta E^1_{it}}{P^1_{i,t-1}}$: IV1	0.35(0.29)	0.2(0.19)	0.47(0.27)	0.35*(0.16)	0.21(0.17)	0.55**(0.19)
$\frac{\Delta E^1_{it}}{P^1_{i,t-1}}$: IV2	0.27(0.27)	0.17(0.21)	0.41(0.25)	0.31(0.18)	0.07(0.13)	0.35(0.18)
$\frac{\Delta E^2_{it}}{P^2_{i,t-1}}$: OLS	0.02(0.02)	0.04(0.03)	0.04(0.02)	0.05*(0.02)	0.01(0.03)	0.02(0.05)
$\frac{\Delta E^2_{it}}{P^2_{i,t-1}}$: IV1	0.08(0.08)	0.04(0.04)	0.13(0.08)	0.08*(0.03)	0.05(0.04)	0.12**(0.05)
$\frac{\Delta E^3_{it}}{P^3_{i,t-1}}$: OLS	0(0.01)	0(0)	0.01(0.01)	0.01(0.01)	0(0.01)	0(0)
$\frac{\Delta E^3_{it}}{P^3_{i,t-1}}$: IV1	0.03(0.03)	0.01(0.01)	0.04(0.03)	0.02(0.01)	0.01(0.01)	0.02*(0.01)

Each cell is the result of a separate regression where the units of observations are migrant sending countries in 2000. Table presents regression results for the impact of the growth in three different groups of emigrants relative to population: all emigrants (All), emigrants with secondary and tertiary education (Secondary and Tertiary), and emigrants with tertiary education (Tertiary) on GDP growth. The methods of estimation are Ordinary Least Squares (OLS) and Instrumental Variable (IV) approach with different instruments constructed for each group of emigrants.'Basic' regressions include only the growth in emigration relative to population as an independent variable while the 'Extended' regressions add to 'Basic' additional control variables: initial dependent variable and human capital, average of government expenditures and trade share in GDP. All control variables in IV regressions are instrumented by their lags to ensure their exogeneity. The numbers in parentheses are heteroskedasticity robust standard errors of the coefficients and (*) indicates significance level at 10, (**) at 5, and (***) at 1 percent.

Table B.9: Decomposing Effects: Human Capital

	All Countries		Non-High Income		Low and Lower-Middle Income	
	Basic	**Extended**	**Basic**	**Extended**	**Basic**	**Extended**
$\frac{\Delta E^1_{it}}{P^1_{i,t-1}}$: OLS	0.05(0.07)	0.11(0.06)	-0.07(0.07)	0.07(0.06)	-0.16(0.09)	0.1(0.07)
$\frac{\Delta E^1_{it}}{P^1_{i,t-1}}$: IV1	-0.2(0.15)	0.19(0.12)	-0.04(0.09)	0.19*(0.09)	-0.07(0.1)	0.14(0.07)
$\frac{\Delta E^1_{it}}{P^1_{i,t-1}}$: IV2	-0.29(0.15)	0.17(0.11)	-0.1(0.09)	0.17(0.09)	-0.14(0.09)	0.12(0.06)
$\frac{\Delta E^2_{it}}{P^2_{i,t-1}}$: OLS	0.02(0.01)	0.01(0.01)	-0.01(0.02)	0(0.01)	-0.05(0.03)	0(0.1)
$\frac{\Delta E^2_{it}}{P^2_{i,t-1}}$: IV1	-0.05(0.04)	0.06(0.04)	-0.01(0.03)	0.05(0.03)	-0.02(0.03)	0.03(0.02)
$\frac{\Delta E^3_{it}}{P^3_{i,t-1}}$: OLS	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
$\frac{\Delta E^3_{it}}{P^3_{i,t-1}}$: IV1	-0.02(0.01)	0.01(0.01)	0(0.01)	0.01(0.01)	0.01(0.01)	0.01(0)

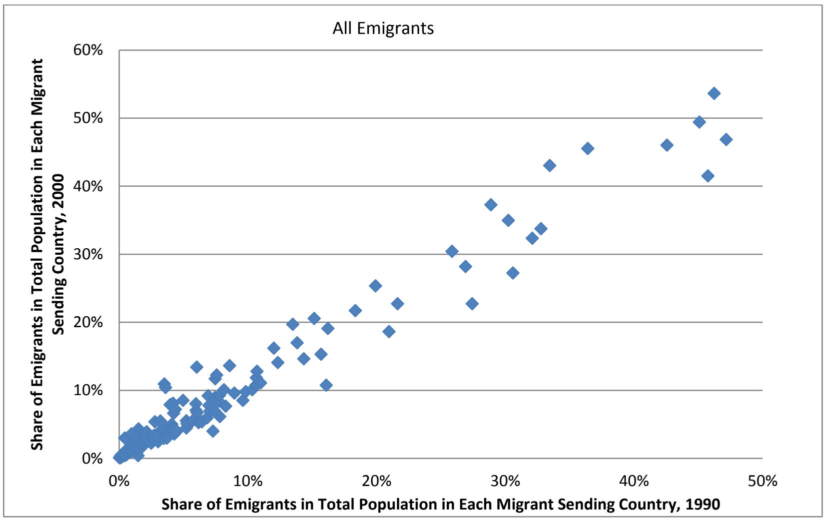
Each cell is the result of a separate regression where the units of observations are migrant sending countries in 2000. Table presents regression results for the impact of the growth in three different groups of emigrants relative to population: all emigrants (All), emigrants with secondary and tertiary education (Secondary and Tertiary), and emigrants with tertiary education (Tertiary) on GDP growth. The methods of estimation are Ordinary Least Squares (OLS) and Instrumental Variable (IV) approach with different instruments constructed for each group of emigrants.'Basic' regressions include only the growth in emigration relative to population as an independent variable while the 'Extended' regressions add to 'Basic' additional control variables: initial dependent variable and human capital, average of government expenditures and trade share in GDP. All control variables in IV regressions are instrumented by their lags to ensure their exogeneity. The numbers in parentheses are heteroskedasticity robust standard errors of the coefficients and (*) indicates significance level at 10, (**) at 5, and (***) at 1 percent.

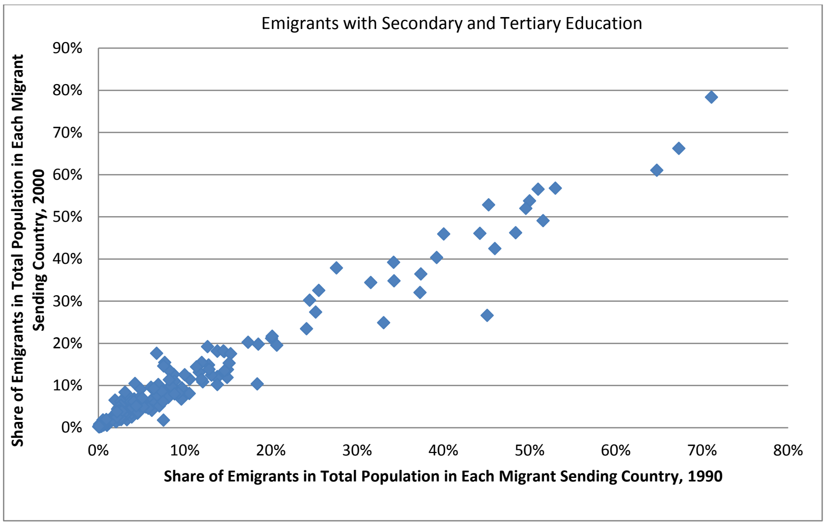
Table B.10: Decomposing Effects: TFP

	All Countries		Non-High Income		Low and Lower-Middle Income	
	Basic	**Extended**	**Basic**	**Extended**	**Basic**	**Extended**
$\frac{\Delta E^1_{it}}{P^1_{i,t-1}}$: OLS	0.64(0.4)	0.77(0.4)	0.98*(0.42)	1.03*0.42)	1.76***(0.5)	1.68*(0.64)
$\frac{\Delta E^1_{it}}{P^1_{i,t-1}}$: IV1	1.28(0.76)	1.144*(0.53)	0.75(0.62)	1.09*(0.44)	1.65**(0.49)	1.76***(0.44)
$\frac{\Delta E^1_{it}}{P^1_{i,t-1}}$: IV2	1.15(0.81)	0.8(0.61)	0.53(0.62)	0.76(0.55)	1.78***(0.4)	1.82***(0.48)
$\frac{\Delta E^2_{it}}{P^2_{i,t-1}}$: OLS	0.06(0.1)	0.11(0.1)	0.11(0.1)	0.12(0.1)	0.42**(0.15)	0.35*(0.16)
$\frac{\Delta E^2_{it}}{P^2_{i,t-1}}$: IV1	0.33(0.23)	0.23(0.15)	0.14(0.18)	0.19(0.13)	0.44**(0.13)	0.43***(0.12)
$\frac{\Delta E^3_{it}}{P^3_{i,t-1}}$: OLS	0.03(0.02)	0.04*(0.02)	0.04(0.02)	0.04*(0.02)	0.061* (0.02)	0.05*(0.02)
$\frac{\Delta E^3_{it}}{P^3_{i,t-1}}$: IV1	0.1(0.07)	0.06(0.03)	0.03(0.04)	0.05(0.02)	0.09***(0.02)	0.085***(0.02)

Each cell is the result of a separate regression where the units of observations are migrant sending countries in 2000. Table presents regression results for the impact of the growth in three different groups of emigrants relative to population: all emigrants (All), emigrants with secondary and tertiary education (Secondary and Tertiary), and emigrants with tertiary education (Tertiary) on GDP growth. The methods of estimation are Ordinary Least Squares (OLS) and Instrumental Variable (IV) approach with different instruments constructed for each group of emigrants.'Basic' regressions include only the growth in emigration relative to population as an independent variable while the 'Extended' regressions add to 'Basic' additional control variables: initial dependent variable and human capital, average of government expenditures and trade share in GDP. All control variables in IV regressions are instrumented by their lags to ensure their exogeneity. The numbers in parentheses are heteroskedasticity robust standard errors of the coefficients and (*) indicates significance level at 10, (**) at 5, and (***) at 1 percent.

Figure B.1: Share of Emigrants in Native Population across Countries by Different Education Groups in 1990 and 2000 (Author's calculations).





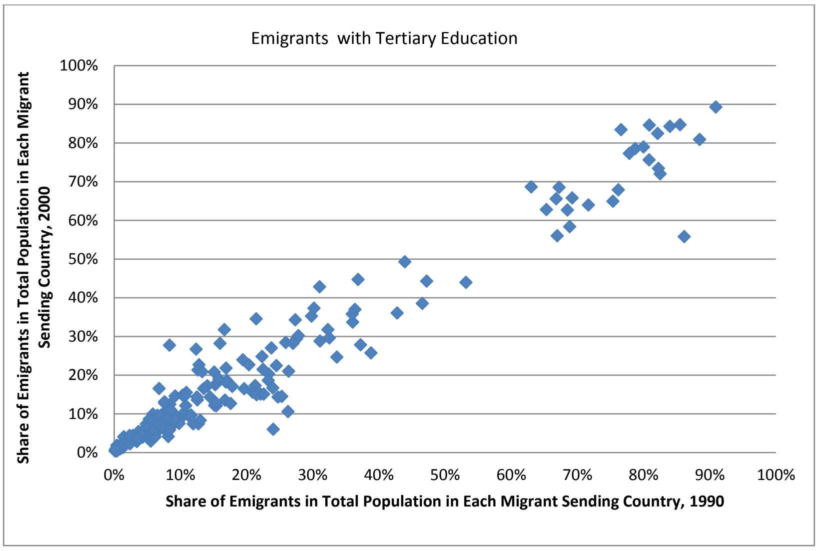
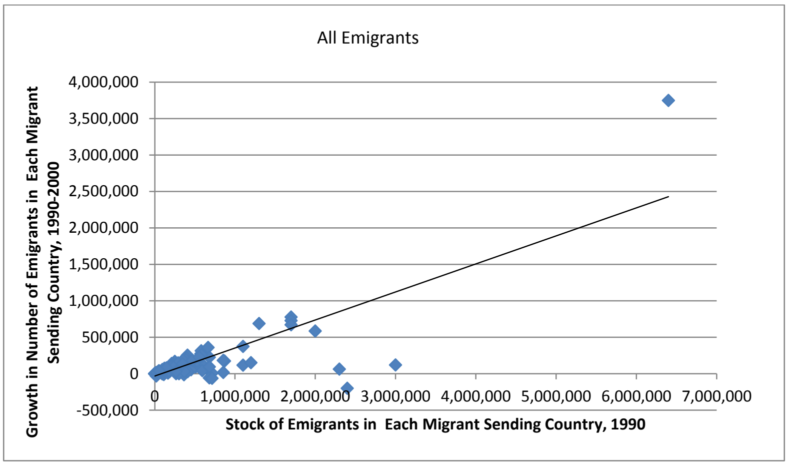
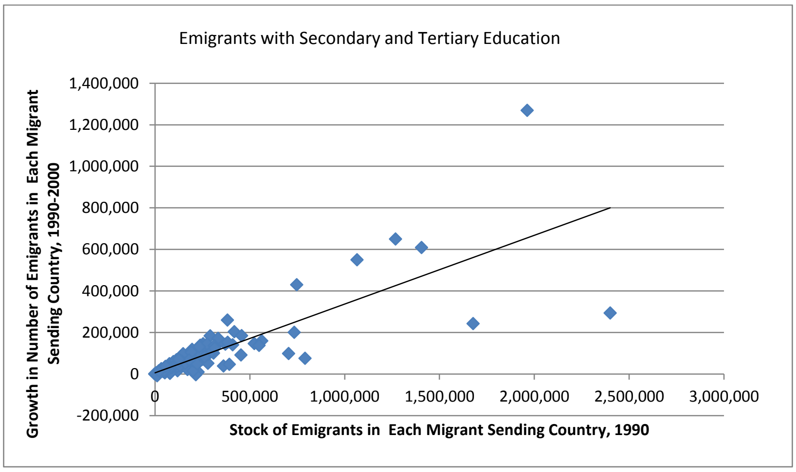
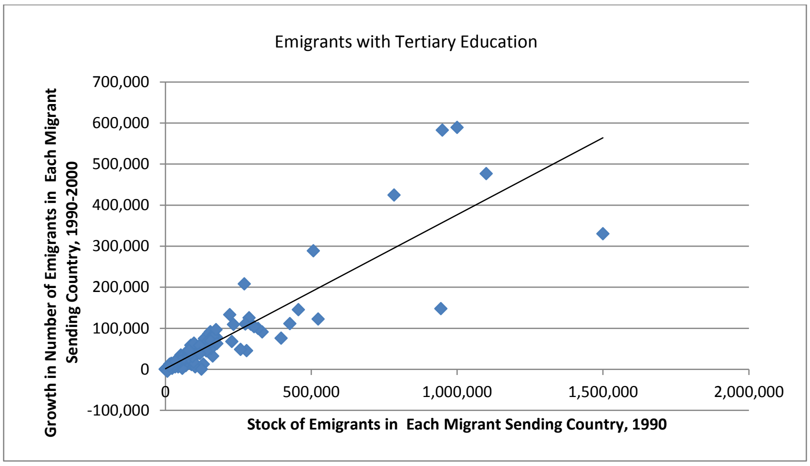

Emigrants with Tertiary Education
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Share of Emigrants in Total Population in Each Migrant Sending Country, 2000
Share of Emigrants in Total Population in Each Migrant Sending Country, 1990

Figure B.2: Stock and Change of Number of Emigrants across Countries by Different Education Groups in 1990 and 2000 (Author's calculations).







Emigrants with Tertiary Education
Growth in Number of Emigrants in Each Migrant Sending Country, 1990-2000
700,000
600,000
500,000
400,000
300,000
200,000
100,000
0
-100,000
0
500,000
1,000,000
1,500,000
2,000,000
Stock of Emigrants in Each Migrant Sending Country, 1990

Figure B.3: Share of Migrants in Total Number of Immigrants by Countries of Origin in the USA in 1990 and 2000 (Author's calculations).

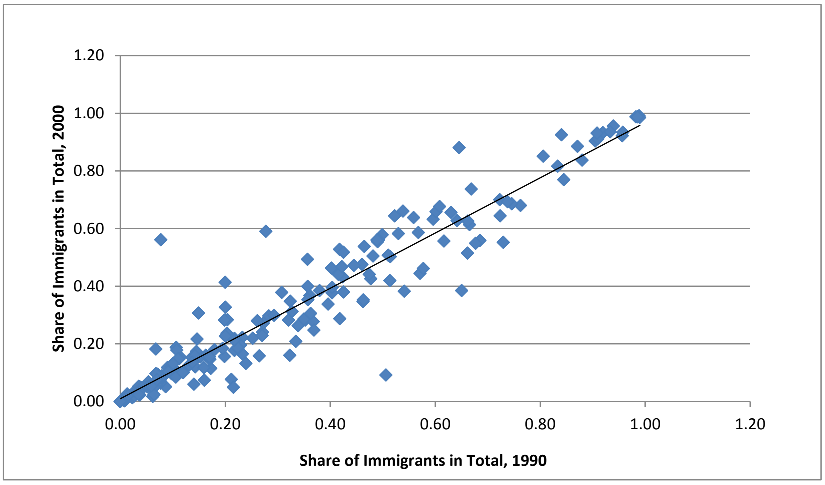
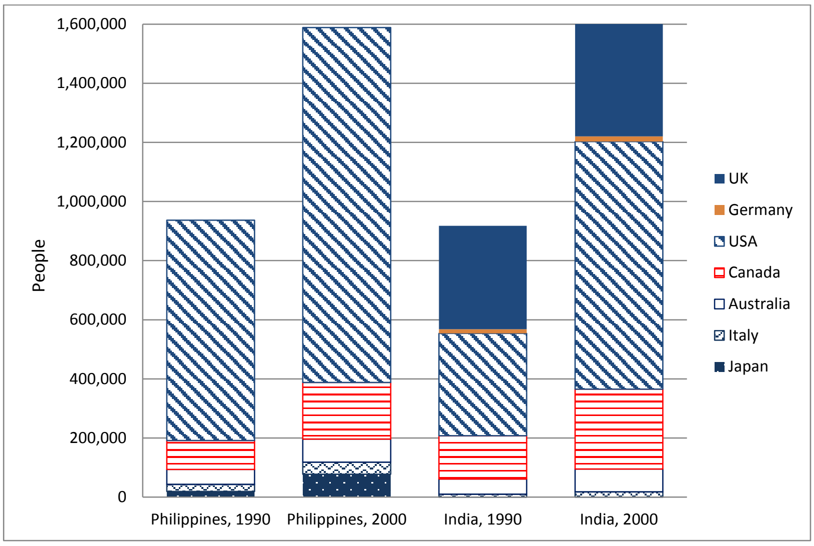


Figure B.4: Total Number of Emigrants in India and Philippines in 1990 and 2000 by Major Destination Countries (Author's calculations).



Chapter 2

Growth Implications of Immigration: Evidence from U.S. Industries

2.1 Introduction

Along with increased flows of capital, goods, and services, international labor mobility has become an inseparable part of globalization, with enormous economic, social and cultural implications in both countries of origin and destination. From 1965 to 2005, the number of international migrants more than doubled to reach three percent of the world's population. During this period, the U.S. has become the country with the largest number of foreign-born inhabitants, hosting about 31 million immigrants in 2000 (Figure $D.1$). The positive impact of immigration on the economies of destination countries goes beyond a pure increase in labor supply. Immigrants bring a different set of work skills, which may complement those of domestic workers and increase productivity (Hunt and Gauthier-Loiselle, 2010). They tend to be young and less risk averse, two traits that are conducive for innovation. By contributing directly to the labor force, immigrants potentially play an important role in generating human capital externalities, as widely recognized in the endogenous growth literature (Friedberg and Hunt, 1995). Finally, immigration fosters competition and encourages specialization of natives towards communication-intensive tasks, leading to increased long-run productivity (Peri and Sparber, 2009). Nevertheless, immigration might have negative effects by crowding out fixed factors of production such as capital.

This paper studies the impact of immigration on the GDP per worker in U.S. industries and its components obtained by production function decomposition: TFP, the capital-output ratio, average hours worked, and skill intensity, defined as a CES function with productivity-weighted high-skill and low-skill labor inputs. The findings show that a one percent increase in the share of immigrants in total employment of an industry leads to 2.24–2.63 percent growth in industry's GDP per worker. This is driven by an increase in TFP (2.08–2.21 percent) and average hours worked (0.23–0.29 percent). Among other components of the production function the skill intensity and capital-output ratio remain unchanged. However, these results are not robust to the inclusion of lagged dependent variables in the regressions.

The empirical literature on immigration has generally focused on its impact on the labor market outcomes of the native population. There are two econometric approaches used to estimate immigration effects: spatial correlation and skill cell. The spatial correlation approach explores the immigration impact on wages and employment of the native population by using data on the geographic distribution of immigrants. These studies, including LaLonde and Topel (1991), Altonji and Card (1990,1991), Borjas, Freman, and Katz (1997), and others, find only a modest impact of immigration on the variables of interest. The skill cell method studies the immigration impact on labor market outcomes using national level data, thus avoiding the bias in estimates possibly caused by local labor market adjustments present in the previous approach. If there is a huge inflow of immigrants to one region, natives will respond by moving to other regions, thus diffusing the impact of immigration beyond the local labor markets. Studies by Borjas (2003, 2006), Ottaviano and Peri (2005, 2006, 2008), and others group labor inputs into skill cells based on education and experience, assuming there is no mobility across these groups. The results of these studies are significant and sizable, but vary considerably depending on model assumptions. Peri (2012) extends the immigration literature by studying the impact of immigrants on aggregate economic behavior using state-level data. In particular, he shows that a one percent increase in state employment due to immigration raises income per worker by 0.5 percent in that state, mostly driven by an improvement in TFP as a result of specialization of low-skill natives in more communication-intensive tasks given their comparative advantage.

Compared to earlier studies, this paper makes two main contributions. First, it uses industry-level data, given significant variation in the utilization of immigrant labor across industries. While there is a substantial variation in the shares of immigrants in states' total employment (Figure $D.2$), a similar trend is observed in the shares of immigrants in total employment across industries (Table $D.2$). The variation across industries is also observed when data are disaggregated into even smaller geographical units at the Metropolitan Statistical Area (MSA) level. Table $D.1$ classifies both MSAs and industries into deciles by the number of immigrants, with decile 1 having the lowest number of immigrants and decile 10 having the largest number of immigrants. Regardless of the size of immigration labor stocks across MSAs, the industry distribution tends to be similar. The estimation of immigration impact on native employment in industries produces insignificant results, indicating that there is no mobility of natives across industries in response to immigration inflows. This is probably due to the high costs of switching industries.

Second, the paper applies consistent and efficient two-step Difference GMM estimation to study the impact of immigration on the variables of interest, applying both internal and external instruments for a share of immigrants in total employment. Internal instruments are simply longer lags of the share of immigrants in total employment. Following Card (2001), external instruments are constructed using allocations of immigrants by 10 groups based on nationality of origin across industries in 1960, inflated by their national growth rates. This instrument imputes a network-driven immigration exogenous to industry-specific developments. The contemporaneous values of external instruments are used to estimate the two-step Difference GMM with external instruments. In addition, this study combines both internal and external instruments in the estimation.

The paper has the following structure. Section 2 introduces the theoretical framework; Section 3 discusses the estimation approach; Section 4 describes the data and construction of the variables; Section 5 presents estimation results; and Section 6 provides conclusions.

2.2 Theoretical Framework

This paper focuses on immigration effects in the U.S. economy, beyond an increase in labor supply, by studying immigrants' impact on GDP per worker and its components obtained by production function decomposition. It assumes that each industry i in the U.S. in year t produces a homogeneous output with a Cobb-Douglas production function:

$$Y_{it} = K_{it}^{\alpha_{it}} \left[X_{it} A_{it} \Phi \left(h_{it} \right) \right]^{1-\alpha_{it}} \tag{2.1}$$

In equation (2.1) Y_{it} is production of a numeraire good, K_{it} is aggregate physical capital, X_{it} is aggregate hours worked, A_{it} is total factor productivity, and α_{it} is the capital share in GDP in industry i and period t. $\Phi(h_{it})$ is an index of skill intensity as defined in the following expression:

$$\Phi(h_{it}) = \left[\left(\beta_{it} h_{it} \right)^{\frac{\sigma-1}{\sigma}} + \left(\left(1 - \beta_{it} \right) \left(1 - h_{it} \right) \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{2.2}$$

In equation (2.2), $h_{it} = H_{it}/X_{it}$ is the share of hours worked by high-skill labor (H_{it}) in total hours worked (X_{it}), and $(1 - h_{it}) = L_{it}/X_{it}$ is the share of hours worked by low-skill labor (L_{it}). High-skill workers are defined as individuals with some college education or more, and low-skill workers are defined as individuals with a high school degree or less. The parameter β_{it} measures a degree of skill bias, with an increase in

β_{it} indicating an increase in productivity of high-skill workers. This functional form assumes that high-skill and low-skill labor are imperfect substitutes with elasticity of substitution σ.

To study per-worker effects of immigration, GDP is divided by total employment:

$$y_{it} = \frac{Y_{it}}{N_{it}} = \left(\frac{K_{it}}{Y_{it}}\right)^{\frac{\alpha_{it}}{1-\alpha_{it}}} \left[x_{it} A_{it} \Phi\left(h_{st}\right)\right] \tag{2.3}$$

Here, y_{it} is GDP per worker, N_{it} is total employment, K_{it}/Y_{it} is the capital-output ratio, and $x_{it} = X_{it}/N_{it}$ is average hours worked in industry i and period t. There are data available for all components of the production function in equation (2.3) with the exception of skill bias β_{it} and factor-neutral productivity A_{it}. These variables are constructed using production function (1) and the condition that the marginal product of each labor input equals its wage, which is also observed in the data. The ratio of average hourly wages of high-skill and low-skill labor can be expressed in terms of their marginal products as in the following equation:

$$\frac{w_{it}^{H}}{w_{it}^{L}} = \left(\frac{\beta_{it}}{1-\beta_{it}}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{h_{it}}{1-h_{it}}\right)^{-\frac{1}{\sigma}} \tag{2.4}$$

In equation (2.4) w_{it}^{H} and w_{it}^{L} are, respectively, average hourly wages of high-skill and low-skill workers. Assuming an elasticity of substitution between high-skill and low-skill labor σ of 1.75, a conventional value proposed by Peri (2012), the skill bias can be constructed using the following expression:

$$\beta_{it} = \frac{(w_{it}^{H})^{\frac{\sigma}{\sigma-1}} (h_{it})^{\frac{1}{\sigma-1}}}{(w_{it}^{H})^{\frac{\sigma}{\sigma-1}} (h_{it})^{\frac{1}{\sigma-1}} + (w_{it}^{L})^{\frac{\sigma}{\sigma-1}} (1-h_{it})^{\frac{1}{\sigma-1}}} \tag{2.5}$$

Further, the factor-neutral productivity A_{it} is computed as a residual. For estimation purposes the production function is expressed in logarithms:

$$\ln y_{it} = \frac{\alpha_{it}}{1-\alpha_{it}} \ln \frac{K_{it}}{Y_{it}} + \ln x_{it} + \ln A_{it} + \ln \Phi(h_{st}) \tag{2.6}$$

The empirical strategy is to estimate the impact of immigration density, i.e., the share of immigrants in total employment, on GDP per worker $(\ln y_{it})$ and each of its components: the capital-output ratio $((\alpha_{it})(1-\alpha_{it})\ln(K_{it}/Y_{it}))$, average hours worked $(\ln x_{it})$, TFP $(\ln A_{it})$, and skill bias $(\ln \Phi(h_{st}))$. Equation (2.7) below serves as a basis for estimation:

$$b_{it} = \delta_t + \gamma_i + \eta_b \frac{N_{it}^F}{N_{it}} + \epsilon_{it} \tag{2.7}$$

Here the independent variable is a share of immigrants in total employment, where N_{it}^F is number of immigrants, N_{it} is total employment in industry i and year t, δ_t and γ_i are respectively a year and industry fixed effects, ϵ_{it} is a zero-mean random shock, and b_{it} is logarithm of GDP per worker $ln(y_{it})$ and each of the components on the right-hand side of equation (2.6). This estimation strategy has the advantage of disentangling the impact of immigration on GDP per worker into different components.

2.3 Estimation Approach

This study applies two-step System and Difference GMM panel data estimation methods developed by Holtz-Eakin, Newy and Rosen (1988), Arellano and Bond (1991), Blundell and Bond (1998) and Arellano and Bover (1995). They provide consistent and efficient estimators when the dataset has a large number of individual observations over a small time horizon. Among other estimators, Ordinary Least Squares coefficients are subject to simultaneity and omitted variable bias. There is a gain in efficiency in applying two-step System and Difference GMM relative to the Instrumental Variable approach due to the use of additional instruments. There are two sets of instruments used in the estimation: internal and external instruments. Internal instruments are simply $(\text{t}-2)$ and longer lags of the share of immigrants in total employment. External instruments are constructed following the methodology proposed by Card (2001), which uses previous settlements of immigrants as an instrument in studying labor market effects of immigration across geographical regions. Immigrants' networks play a key role in their location choices, as having individuals from the same countries of origin and, therefore, having access to information, substantially reduces migration costs, and drives migrants to the places with higher concentrations of immigrants. This network-driven immigration is exogenous to local labor market developments and can be used as an instrument.

Data investigation shows that this approach is also valid for industry-level analysis. Using data from the public use micro-data samples (IPUMS) of the U.S. Decennial Census and the American Community Survey (ACS), Figures $D.3$ and $D.4$ demonstrate both cross-state and cross-industry distribution of immigrants over the period of 1960–2005. Each point on the left-hand side graphs is the share of foreign-born workers in the total employment of each state, while each point on the right-hand panels is the share of foreign-born workers in the total employment of each industry. These visual comparisons of the distribution of foreign-born employees across states and industries over time indicate that immigrants choose not only to settle in the areas with a prior high concentration of immigrants, but they also tend to be employed in the industries with an existing sizable presence of immigrants. Moreover, the trend is even stronger for industry level data than for state level data, thus showing that networks are also important for immigrants' employment allocations.

This approach classifies immigrants into 10 following groups according to their birthplace or nationality of origin based on U.S. Census data: (1) Mexico, (2) rest of Latin America (including Central America, Caribbean, and South America), (3) Canada, Australia and New Zealand, (4) Western Europe, (5) Central and Eastern Europe and Republics of the Former Soviet Union, (6) China (including Hong Kong, Macau, Mongolia and Taiwan), (7) India (including Bangladesh, Bhutan, Burma, Pakistan and Sri Lanka), (8) rest of Asia, (9) Africa and (10) others. For each nationality of origin n the total number of immigrants in each industry i in 1960, $Pop_{n,i,1960}$, is constructed. Next, national growth rates for each nationality of origin n in the U.S. in

1970, 1980, 1990, 2000, and 2005 relative to 1960 are computed:

$$G_{n,t-1960} = \frac{Pop_{n,t} - Pop_{n,1960}}{Pop_{n,1960}} \tag{2.8}$$

Here $Pop_{n,t}$ is the number of immigrants with nationality of origin n in year t. These national growth rates are applied to the number of immigrants from each nationality of origin n in each industry i in 1960 in order to impute the number of immigrants by the nationalities of origin in the subsequent decades across industries. Thus, the imputed number of immigrants from the nationality of origin n in industry i and year t is:

$$\widehat{Pop}_{n,i,t} = Pop_{n,i,1960} * [1 + G_{n,t-1960}] \tag{2.9}$$

The imputed total number of immigrants in each industry i is obtained by summing across the nationalities of origin:

$$\widehat{Pop}_{Fi,t} = \sum_{n} \widehat{Pop}_{n,i,t} \tag{2.10}$$

Finally, the instrument for the share of immigrants in the total employment for each industry is constructed in the following way:

$$Instrument = \frac{\widehat{Pop}_{Fi,t}}{\widehat{Pop}_{Fi,t} + Pop_{USi,t}} \tag{2.11}$$

where $Pop_{USi,t}$ is an actual number of working natives in industry i and year t.

The validity of this external instrument is also consistent with the history of U.S. immigration. A closer look at U.S. immigration policies shows that during this period of time, immigration was primarily based on family ties or kinship with the U.S. citizen or legal immigrant rather than industry-specific labor demand. According to Jasso and Rosenzweig (1990) the history of immigration legislation in the 20th century suggests that there were five principal aims of immigration law: (i) to avoid large increases in the foreign-born population, (ii) to avoid substantial shifts in the country-of-origin composition of the foreign-born, (iii) to facilitate the unification of immediate family relatives regardless of their place of birth, (iv) to facilitate the acquisition of scarce labor skills by U.S. employers, and (v) to provide a refuge for displaced persons, chiefly those threatened by foreign governments. Historically, the U.S. enacted two main immigration laws: the Immigration Act of 1924, which constrained entry into the U.S. for the first time, and its later modification in 1965. The 1924 Immigration Act set limits on immigration from the Eastern Hemisphere, restricting it to two percent of national origin of 1890 foreign-born, with a maximum of 164,000 people. From 1924 to 1965–1968, potential immigrants with scarce skills in agriculture were accorded the highest preference among numerically restricted immigrants. As skilled agricultural immigration can generate a simultaneity bias in the estimates due to the fact that skilled immigration can cause industry growth and the growth of the industry can trigger more immigration, agriculture is excluded from the analysis to avoid this endogeneity bias. The Immigration and Nationality Act of 1965 abolished national-origin quotas and established uniform per-country limits of $20,000$ and a preference category system with an overall ceiling of $170,000$ for the Eastern Hemisphere. By 1968, the annual limitation from the Western Hemisphere was set at $120,000$ immigrants, with visas available on a first-come, first-served basis. At the same time there were certain categories exempt from the limitation, such as immediate relatives of U.S. citizens, refugees and asylum seekers adjusting to permanent residence and special immigrants such as certain foreign medical graduates. According to Jasso and Rosenzweig (1990), kinship with a U.S. citizen or legal immigrant during 1965–1990 was the principal route of immigration to the United States. No more than 20 percent of numerically limited visas were allocated to occupation-based applicants and their families, and less than four percent of all immigrants were screened with respect to labor market criteria. In addition certain foreign medical graduates included in the non-quota immigrants' pool accounted for only $3,000$ individuals or about 0.07 percent of total number of immigrants. Thus, the examination of immigration flows over the period of 1924–1990 indicates that an insignificant share of labor mobility was driven by industry-specific labor demand in the U.S. with the exception of agriculture. Therefore, immigration to the U.S. was mostly supply-driven, exogenous to industry-specific conditions, and networks tended to play a major role in immigrants' distribution across industries.

In addition, this paper combines both internal and external instruments in estimations to study the impact of immigration on the variables of interest.

2.4 Data Description

There are three sources of data used in this study. First, the aggregate industry-specific variables, including GDP and capital stock, are obtained from the sectoral input-output database developed by Dale W. Jorgenson, which covers 35 sectors at the two-digit Standard Industrial Classification (SIC) level from 1960 to 2005. It provides quantities and prices of industries' output, capital stock and production inputs. The industry-specific capital shares used in the production function decomposition are also constructed using this dataset.

Next, the micro data on the number of working immigrants and natives across industries and their characteristics, including education, hours worked and wages, are computed from the public use micro-data samples (IPUMS) of the U.S. Decennial Census and the American Community Survey (ACS). This dataset combines the 1960 one percent Sample, the 1970 one percent State Sample Form 1, the 1980 one percent Sample, the 1990 one percent Sample, the 2000 one percent Sample from the U.S. Census data, and the 2005 ACS. For the weighted regression analysis the 1950 one percent sample is added to compute total employment by industries in the pre-sample period. The dataset includes only individuals of age 17 and older who don't live in Group Quarters, and worked in the previous year reporting valid salary income with an experience of more than one and less than 40 years. The workers are classified into two educational groups: low-skill labor and high-skill labor. Low-skill labor includes individuals with a high-school degree or less, and high-skill labor includes individuals with some college education or more. Immigrants are defined as naturalized citizens or non-citizens following the conventional approach in the literature. To compute hours worked used as production inputs for high-skill and low-skill labor, hours worked in a week are multiplied by weeks worked in a year. This individual labor supply is further aggregated by education and industry groups using "perwt" frequency weights for individuals. The hourly wages of individuals are constructed as yearly wages divided by the total hours worked, which then are averaged by education and industry groups using total hours worked by individuals as weights.

The allocation of immigrants across industries in 2005 is shown in Table $D.2$ of the Data Appendix section. Top five industries which hire the highest number of immigrants are Apparel, Agriculture, Leather, Food and kindred products, and Miscellaneous manufacturing. The share of foreign-born population in the total employment of these industries accounts for more than 25 percent. Instead, Utilities, Tobacco, Non-metallic mining, Metal mining, and Coal mining are industries with the least shares of immigrants in total employment ranging from 1 to 9 percent.

Finally, data from the O*NET Abilities Survey conducted by the U.S. Department of Labor contain information on employee abilities used to compute the task specialization variable for each industry. This variable is constructed using manual and communication tasks. The manual task is an average of the following 19 abilities or tasks: arm-hand steadiness level, manual dexterity level, finger dexterity level, control precision level, multilimb coordination level, response orientation level, rate control level, reaction time level, wrist-finger speed level, speed of limb movement level, static strength level, explosive strength level, dynamic strength level, trunk strength level, stamina level, extent flexibility level, dynamic flexibility level, gross body coordination level, and gross body equilibrium level. The communication task is computed as an average of four tasks: oral comprehension importance, written comprehension importance, oral expression importance, and written expression importance. These tasks are assigned to individuals in 1960–2000 Census and the 2005 American Community Survey data from IPUMS based on their occupation.

2.5 Estimation Results

2.5.1 Model without Lagged Dependent Variable

The empirical approach is to estimate Equation (2.7) for the GDP per worker and its components as shown in Equation (2.6): capital-output ratio, average hours worked, TFP, and skill intensity and its components: skill bias and hours worked by high-skill labor. The explanatory variables are the share of immigrants in total employment of industries and time and industry fixed effects to control for time and industry specific shocks. The estimation approach is standardized to two-step Difference GMM since in most cases there is no first-order serial correlation in the error term. This implies a random walk which means the two-step System GMM estimator is not consistent. The two-step Difference GMM addresses the endogeneity issues possibly caused by industry-specific shocks since they take the first-differences of the variables and use the lagged values of the level variables as instruments. Three types of instruments are used in these estimations: internal, external, and all instruments, which combine both internal and external instruments. The internal instruments are lagged values of the share of immigrants in industry's total employment taken with $(t-2; t-3)$ lags where t is a decade. The external instruments are constructed using the distribution of immigrants across industries in 1960 inflated by their national rate of decade growths. To check the robustness of the results, the $(t-1)$ lagged values of the dependent variables are added to the regressions. The lagged values of the dependent variable are, in turn, instrumented by their $(t-2; t-3)$ lags. If there is second-order serial correlation in errors as tested by Arellano–Bond test, then $(t-3)$ lags are used. Also, the contemporaneous values of external instruments are applied. This paper uses both unweighted and weighted estimations. Using weights follows Peri (2012). Two weights are used for estimations: industries' employment in 1950 and average employment of industries over the period of 1960–2005.

Each cell in Table $D.3$ shows the estimated coefficients of immigration shares in total employment (η_b) of a separate regression with heteroscedasticity-robust standard errors. These regressions are not weighted and the units of observation are U.S. industries in each census year for a period of 1960–2000 and 2005. Table $D.3$ reports the two-step Difference GMM estimation results with internal, external, and all instruments, which combine both internal and external instruments. The Hansen test of overidentifying restrictions for all estimates indicates that both internal and external instruments are independent of the disturbance process, hence, can be used as valid instruments.

Table $D.3$ also includes regression results of equation (2.7) for the logarithm of native employment as a dependent variable to address some concerns raised in the literature. The existing empirical work focusing on growth effects of immigration is limited, and one of the recent studies by Peri (2012) uses state-level data on immigration and aggregate economic variables to test whether foreign-born workers simply increase labor supply or alter the behavior of other economic variables. However, results based on regional distribution of immigrants might be potentially misleading, as in a highly mobile U.S. labor market there would be an adequate response from natives to immigration inflows. The domestic workers in states with a high concentration of immigrants would choose to move to other states with more favorable local labor market conditions, or workers in other states would prefer not to move to states with a large number of immigrants. Therefore, immigration effects might spread across states. An extended literature in labor economics studying the impact of immigration on labor market outcomes encountered these problems in using regional data, which generated insignificant and small effects of immigration. More recent studies generally focus on national level data, classifying the labor inputs by education and experience groups, which minimizes the mobility of domestic labor across these groups in response to immigration. This paper tests if there are any implications on the native employment of industries in response to immigration. Table $D.3$ reports estimation results for the impact of immigration on the native employment. Regression results with different instruments and for different specifications generally indicate that the immigrants crowd out natives.

The regression results in Table $D.3$ indicate that immigrants' density, i.e. a share of foreign-born workers in total employment of an industry, positively affects the GDP per worker in the regressions with external and all instruments at the 10 percent significance level. An increase in the share of immigrants in industry's total employment by one percentage point raises industry's GDP per worker by about 2.24–2.63 percent in these specifications. The production function decomposition of the GDP per worker helps understand the channels of immigration impact. According to the results in Table $D.3$, there is no change in capital intensity or the capital-output ratio in response to immigrants' inflows in the regressions with internal, external, and all instruments. This indicates that the capital-output ratio adjusts to its long-term trend with an increase in labor. Average hours worked increase by $0.23 - 0.29$ percent in all specifications. The results are significant at the10 percent level when using internal and external instruments separately and at the 1 percent level when these instruments are used in combination. The estimation results with different instruments demonstrate that immigration has no impact on the skill intensity. Among the components of the skill intensity, however, productivity of high-skill labor or skill bias increases when applying external and all instruments with a magnitude of $2.03 - 2.48$ at the five percent significance level. The estimation results indicate that hours worked by high-skill labor increases in response to immigration by 2.41–2.82 percent. Finally, the major driving force of an increase in the GDP per worker is factor neutral productivity, or TFP, which increases by 2.08–2.21 percent for specifications with external and all instruments at 1 percent significance level. Considering that the average share of immigration in the total employment of an industry was 0.15 in 2000, then the GDP and TFP increases would be respectively at most 0.4 and 0.3 percent for a decade-long change.

Table $D.4$ reports the regression results for the weighted two-step Difference GMM estimations without the lagged dependent variable. While the validity of the instruments in the weighted regressions is questionable, this approach has been used in several papers such as Peri (2012). Weighting regressions by industry's employment might generate a correlation between the errors and the instrument which is an imputed share of immigrants in industry's employment. Two time-invariant weights have been chosen: average employment of industries over the period of $1960 - 2005$ and employment of industries in 1950. Table $D.4$ presents estimation results only with the average employment of industries used as weights as the results are similar in terms of signs, significance, and magnitudes when employment of industries in 1950 is applied. The estimates yield similar results for GDP per worker and TFP as in the non-weighted case but with improved significance and larger magnitudes of the coefficients. The GDP per worker increases by nearly 5.8 percent in response to immigration change mostly driven by positive changes in the TFP with a magnitude of $5.24 - 5.53$ percent when external and all instruments are used. These results are significant at the 10 percent level for the GDP per worker and the five percent level for TFP. The capital-output ratio and the skill intensity remain unchanged as in non-weighted regressions despite an increase in the components of the latter. The immigration impact on the average hours worked is 0.27 at the five percent significance level when internal instruments are used and the coefficient loses its significance when external and all instruments are applied.

2.5.2 Model with Lagged Dependent Variable

The robustness of the results are further tested with an inclusion of the lagged dependent variables in the regressions. The unweighted estimation results are reported in Table $D.5$. Tables $D.5$ also provides estimation results for regressions with all instruments augmented with an interaction term of the lagged dependent variable and a dummy variable for year 2005, as the number of years between the last two time periods in data is five compared to 10 for other time points. Inclusion of the lagged dependent variable qualitatively alters the significance of the results for the impact of immigration. In all estimations the lagged dependent variables are significant with an exception of the capital-output ratio for all specifications and hours worked by high-skill labor when using external instruments.

The coefficient of the immigration impact on the GDP per worker retains its significance and sign only in the specification with all instruments while is insignificant in all other specifications. Instead, the results show that there is no impact of immigration on TFP. Capital-output ratio and skill intensity remain unaffected with changes in immigration as in the model without the lagged dependent variable. Among the components of the skill intensity, the impact of immigration on the skill bias and hours worked by high-skill labor is similar to the model without the lagged dependent variable. The skill bias increases with immigration by $0.51 - 1.2$ percent when external and all instruments are applied and is unchanged in specifications with internal instruments and all instruments with interaction term. Hours worked by high-skill labor increase across all specifications with a coefficient in the range of $1.15 - 1.65$.

Following the literature, Table $D.6$ reports the weighted regression results with time-invariant average employment of industries used as weights. Estimation of the weighted two-step Difference GMM produce results varying from the unweighted estimates. This is the only specification when the impact of immigration on the native employment is insignificant when applying different instruments. The GDP per worker, TFP, skill intensity, skill bias, and hours worked by high-skill labor remain unchanged in response to immigration. Instead, the capital-output ratio increases by $0.23 - 0.25$ percent at the 10 percent significance level when all instruments and all instruments with interaction term are applied. The average hours increase by 0.71 percent when all instruments are applied. These variables remain unaffected by in all other specifications.

2.5.3 Different Elasticities of Substitution between high-skill and Low-skill Labor

Among the variables of interest, only the TFP and skill intensity are constructed using theoretical assumptions, while the rest of the variables are directly observed in the data. To test the robustness of the results, TFP and skill intensity are computed for different values of the elasticity of substitution between high-skill and low-skill labor σ. Table $D.7$ reports the estimation results with the lagged dependent variables and shows that immigration has no impact on the TFP when different values of σ are used. Instead, the significance and signs of the coefficients on the impact of the immigration on the skill intensity change as σ increases. There is no impact on the skill intensity when σ is low and equals 1.5, 1.75, or 2. The skill intensity increases by 0.23–0.40 when σ equals five and internal, all, and all instruments with interaction terms are used. The only insignificant estimator is when external instruments are applied.

2.5.4 Task Specialization

One of the explanations for increased factor-neutral productivity across the states suggested by Peri (2012) based on Peri and Sparber (2009) is a task specialization of natives in response to immigration flows. In the low-skilled group, if there is a high inflow of foreign-born workers, natives will switch to communication-intensive tasks exploiting their comparative advantage of language skills and familiarity with networks and leave manual intensive tasks to immigrants. Following Peri and Sparber (2009) this paper constructs the task specialization variable for industries and uses it as a control variable in the regressions. The O*NET Abilities Survey conducted by the U.S. Department of Labor provides information on the importance of 52 employee abilities or tasks expressed as numerical values for different occupations. These can be used to compute natives' aggregate supply of communication and manual tasks by education and industry groups as follows. First, this study merges O*NET Abilities Survey with 2000 Census data and assigns each individual a specific percentile score from 0 to 1 for each ability or task according to his occupation and relative importance of the task among all workers in 2000. For example, a percentile task score of 0.05 for the occupation indicates that only five percent of workers in the U.S. were supplying that task less intensely in 2000. Next, the paper computes individuals' manual task supply as an average of percentile scores for 19 abilities or tasks. To construct supply of communication tasks it takes the average of scores for the following four abilities or tasks: oral comprehension importance, written comprehension importance, oral expression importance, and written expression importance. These average scores are assigned to individuals in $1960-2000$ Census and 2005 American Community Survey data taken from IPUMS based on their occupations. Demographic characteristics such as experience, education, gender, and race are used in the first stage regressions to clean these individual task supply variables as they might be correlated with immigration. Industry averages of individuals' cleaned supply of communication and manual tasks are computed using their personal weights and hours worked as weights. Finally, the relative task supply, expressed as a logarithm of the ratio of the average supplies of manual and communication tasks, is used to control for the task specialization in the regressions. The paper tests whether the results are changed with an inclusion of this variable. The instruments used for the share of immigrants in total employment are also valid for the task specialization variable. As shown in Table $D.8$, the coefficients on the impact of immigrants on the TFP remain insignificant when controlling for task specialization. The only two significant coefficients for the skill bias in specifications with external and all instruments lose their significance when task specialization is included in the regressions.

2.6 Conclusion

This paper discusses the growth implications of immigration in the U.S. by estimating its impact on the GDP per worker, TFP, the capital-output ratio, average hours worked, and skill intensity, defined as a CES function with productivity weighted high-skill and low-skill labor inputs. It applies the two-step Difference and System GMM to estimate the impact of immigrants' share in total employment on the variables of interest. The paper uses internal instruments, which are lagged values of immigrants' share in total employment; external instruments constructed based on immigrants' distribution across industries in 1960 and their national growth rates; and all instruments combined. The estimation results show some positive impact of the immigration as a share of total employment on the GDP per worker, generally driven by improvements in the factor-neutral productivity. The consistent and efficient two-step Difference GMM estimators indicate that a one percent increase in the share of foreign-born workers in total employment increases GDP per worker by nearly $2.24 - 2.63$ percent. The decomposition of the production function suggests that the main channel of the impact is TFP, which increases by 2.08–2.21 percent. In addition, average hours worked grow by 0.23–0.29 percent across different specifications. The skill intensity and the capital-output ratio remain unchanged. However, these results are not robust to the inclusion of the lagged dependent variables.

Bibliography

[1] Altonji, J. and D. Card (1991), "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives", in Abowd, J. and R. Freeman, eds., Immigration, Trade and the Labor Market. Chicago: University of Chicago Press, 201-234.

[2] Andersson, F., J. Haltiwanger and S. Sanders (2006), mimeo, University of Maryland.

[3] Autor, D., L. Katz, and A. Krueger (1998), "Computing Inequality: Have Computers Changed the Labor Market?", The Quarterly Journal of Economics 113 (4): 1169-1213.

[4] Borjas, G. (2004), "Increasing the Supply of Labor Through Immigration Measuring the Impact on Native-born Workers," Center for Immigration Studies Backgrounder.

[5] Borjas, G. (2003), "The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," Quarterly Journal of Economics 118 (4): 1335-1374.

[6] Borjas, G. (1999), Heaven's Door. Princeton: Princeton University Press.

[7] Borjas, G. (1995), "Assimilation and Changes in Cohort Quality Revisited: What Happened to Immigrant Earnings in the 1980s?" Journal of Labor Economics 13(2): 201-245.

[8] Card, D. (2001), "Immigrant Inflows, Native Outflows, and the Local Market Impacts of Higher Immigration", Journal of Labor Economics 19: 22-64.

[9] Freidberg, R. and J. Hunt (1995), "The Impact of Immigrants on Host Country Wages, Employment and Growth", Journal of Economic Perspectives 9: 23-44.

[10] Hansen, G. and M. Slaughter (2002), "Labor Market Adjustments in Open Economies: Evidence from US States" Journal of International Economics 57 (1):3-29.

[11] Hovhannisyan, S. (2012), "Labor Market and Growth Implications of Emigration: Cross-Country Evidence", The World Bank: Background Paper for the World Development Report 2013.

[12] Hunt, J. and M. Gauthier-Loiselle (2010), "How Much Does Immigration Boost Innovation?", American Economic Journal: Macroeconomics, American Economic Association, 2(2): 31-56.

[13] Jorgenson, D.W. (2007), "35 Sector KLEM", http://hdl.handle.net/1902.1/10684 UNF:3:TqM00zRqsatX2q/teT253Q==V1 [Version]

[14] Jasso, G. and M.R. Rosenzweig (1990), "The New Chosen People: Immigrants in the United States", The Russel Sage Foundation.

[15] LaLonde, R. and R. Topel (1991), "Labor Market Adjustments to Increase Immigration." In Abowd, J., and R. Freeman, eds., Immigration, Trade and the Labor Market. Chicago: University of Chicago Press: 167-200.

[16] Lewis, E. (2006), ”Immigration, Skill Mix, and the Choice of Technique”, mimeo, Dartmouth College, and Federal Reserve Bank of Philadelphia Working Paper 05-08.

[17] Lewis, E. (2005), ”The Impact of Immigration on New Technology Adoption in U.S. Manufacturing”, mimeo, Federal Reserve Bank of Philadelphia.

[18] Lewis, E. (2003), ”Local, Open Economies Within the US: How Do Industries Respond to Immigration?”, Federal Reserve Bank of Philadelphia Working Paper No. 04-1.

[19] Ottaviano, G.I.P. and G. Peri Giovanni (2008), ”Immigration and National Wages: Clarifying the Theory and the Empirics” NBER Working Paper 14188.

[20] Ottaviano, G. and G. Peri (2006), ”Rethinking the Effects of Immigration on Wages”, NBER Working Paper 12497.

[21] Ottaviano, G. and G. Peri (2006), ”The Economic Value of Cultural Diversity: Evidence from U.S. cities”, Journal of Economic Geography, Vol. 6(1): 9-44.

[22] Peri, G. (2012), ”The Effect of Immigration on Productivity: Evidence from U.S. States”, The Review of Economics and Statistics, Vol. 94(1): 348-358.

[23] Peri, G. and C. Sparber (2009), ”Task Specialization, Immigration and Wages”, American Economic Journal: Applied Economics (1):3.

[24] Peri, G. and C. Sparber (2007), ”Comparative Advantage and Gains from Immigration”, mimeo, University of California, Davis. Singer, A. (2004), ”The Rise of New Immigrant Gateways”, Brookings Institution Center on Urban and Metropolitan Policy.

[25] Singer, A. (2004), ”The Rise of New Immigrant Gateways”, Living Cities Series, Brookings Institution Center on Urban and Metropolitan Policy.

Appendix C

Data Appendix

Jorgenson's dataset contains information on quantities and prices of four inputs, including capital, labor, energy, and material, and output for 35 industries. For the analysis Motor vehicles and Transportation equipment and ordnance industries are combined into one industry - Motor vehicles, transportation equipment and ordnance, and Electric utilities and Gas utilities are combined into one industry - Utilities, since there is no direct mapping between 1987 SIC codes and Jorgenson's industry classifications. Table $D.2$ shows the classification of industries ranked by the share of immigrants in total employment of industry in descending order. Agriculture, government and mining sectors are dropped from the analysis.

The labor market variables are constructed using the 1960 one percent Sample, the 1970 one percent State Sample Form 1, the 1980 one percent Sample, the 1990 one percent Sample, the 2000 one percent Sample from the U.S. Census data, and 2005 ACS. The sample for the analysis is selected using the following criteria:

- include only individuals of age 17 and older corresponding to the age 16 and older in the previous year, since the U.S Bureau of Labor Statistics (BLS) classifies people of working age as those 16 and older, and questions on work variables in the Census relate to the previous year: the variable "age" is higher or equal to 17;
- include only individuals who don't live in "Group Quarters", which according to

the Census definition are largely institutions and other group living arrangements such as rooming houses and military barracks: the variable "gq" is not equal to 0, 3, or 4;

- include only individuals who worked last year: "wkswork2" variable is not equal to 0 for 1960 and 1970, and "wkswork1" variable is not equal to 0 for the years after 1970;
- include only individuals who reported valid salary income: "incwage" variable is not equal to 0 or to 999,999;
- include only individuals with experience more than 1 and less than 40 years. The experience variable is constructed as "(age - time first worked)", where the "time first worked" is 16 for workers with education equal or below Grade 9: variable "educ" is less or equal to 3; 17 for workers with education of Grade 10: "educ"=4'; 18 for workers with education of Grade 11: "educ"=5; 19 for high-school graduates: "educ"=6; 20 for workers with 1 year of college education: "educ"=7; 21 for workers with 2 years of college education: "educ"=8; 22 for workers with 3 years of college education: "educ"=9; and 23 for college graduates: "educ" is larger or equal to 10;
- include only individuals who are not self-employed and unpaid family workers: variable "classwkd" is between 20 and 28.

Below are constructed variables used in the analysis:

- Labor: individuals are classified into 2 educational groups: (1) workers with high-school degree or less ("educ" is less than or equal to 6), and (2) workers with some college education or more ("educ" is larger than 6).
- Immigrants: individuals are considered immigrants if they are not citizens or are naturalized citizens ("bpld" is larger than 15,000, except of 90011 and 90021 for 1960 and "citizen"=2 or 3 for the years after 1960).

- Weeks worked in a year by an individual: since 1980 ”wkswkd1” variable reports the exact number of the weeks worked. For 1960 and 1970 the variable ”wkswork2” is used, which defines weeks worked in intervals. For each interval the median value is used to compute the weeks worked in the previous year for individuals: 6.5 weeks if wkswork2=1, 20 weeks if wkswork2=2, 33 weeks if wkswork2=3, 43.5 weeks if wkswork2=4, 48.5 weeks if wkswork2=5, and 51 weeks if wkswork2=6.

- Hours worked in a week by an individual: ”hrswork2” variable is used for 1960 and 1970 and ”uhrswork” after 1970.

- Hours worked in a year by an individual: this variable is constructed as hours worked in a week multiplied by weeks worked in a year. This is a measure of labor supply by an individual.

- Yearly wages: to protect the confidentiality of respondents, the U.S. Census Bureau ”top codes” values at the extreme upper end of many variable distributions. For the income after 1980, the amounts higher than the topcode are reported as the state median of all values exceeding the topcode. To make the numbers comparable across years the topcodes for yearly wages are multiplied by coefficients 1.5 as used in Peri (2009) for 1960, 1970 and 1980. Wages are also adjusted for the price changes. To express wages in constant prices of 1996 - a year, relative to which prices are normalized in Jorgenson’s dataset, CPI inflation calculator from the BLS website is used. It applies the average Consumer Price Index for a given calendar year, which represents changes in the prices of all goods and services purchased for consumption by urban households.

- Hourly wage of an individual: this is constructed as yearly wage divided by the total hours worked by an individual. Average hourly wages by industries and education groups are constructed using the total hours worked by individuals as weights.

- The task specialization variable is constructed using manual and communication tasks taken from O*NET Abilities Survey. Manual task is an average of the following 19 abilities or tasks: arm-hand steadiness level, manual dexterity level, finger dexterity level, control precision level, multilimb coordination level, response orientation level, rate control level, reaction time level, wrist-finger speed level, speed of limb movement level, static strength level, explosive strength level, dynamic strength level, trunk strength level, stamina level, extent flexibility level, dynamic flexibility level, gross body coordination level, and gross body equilibrium level using $occ1990$ occupational variable. The communication task is computed as an average of four tasks: oral comprehension importance, written comprehension importance, oral expression importance, and written expression importance. These tasks are assigned to individuals in $1960 - 2000$ Census and 2005 American Community Survey data taken from IPUMS based on their occupation using $occ1990$ occupational variable from IPUMS. Also, $Hispan$ and $Race$ variables are used to identify the race of individuals.

The above mentioned variables are computed for each education and industry groups with ”perwt” variable used as weights to obtain aggregate or average levels.

Appendix D

Tables and Figures

Table D.1: Immigrant Distribution by Deciles across Industries and Metropolitan Statistical Areas (MSAs)

MSA/Industry[1]	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	2	2	3	3	3	6	8	10	12	52
2	1	1	2	3	4	3	7	9	14	57
3	1	2	2	4	2	4	7	12	15	51
4	1	1	2	3	3	3	8	9	14	56
5	1	1	2	3	3	5	8	10	15	52
6	0	1	1	3	4	4	7	9	14	55
7	0	1	1	2	3	5	7	9	14	57
8	1	1	1	2	3	5	7	9	15	58
9	1	1	1	2	3	5	7	10	15	57
10	0	1	1	2	3	4	7	11	18	53

[1] This table classifies both MSAs and industries into deciles by the number of immigrants, with decile 1 having the lowest number of immigrants and decile 10 having the largest number of immigrants. The rows show the distribution of immigrants across deciles of industries in each decile of MSAs.

Table D.2: Industries Ranked by Share of Immigrants in Total Employment, 2005

Share	**Industry**	**Share**	**Industry**
0.43	Apparel	0.15	Lumber and wood products
0.39	Agriculture	0.14	Personal and business services
0.29	Leather	0.14	Transportation and warehousing
0.27	Food and kindred products	0.13	Finance Insurance and Real Estate
0.25	Misc. manufacturing	0.13	Printing, publishing and allied
0.24	Electrical machinery	0.12	Primary metal
0.24	Construction	0.12	Communications
0.23	Furniture and fixtures	0.12	Motor vehicles and transp. equipment ordnance
0.21	Instruments	0.11	Paper and allied
0.19	Textile mill products	0.10	Petroleum and coal products
0.18	Stone, clay, glass	0.10	Oil and gas extraction
0.17	Chemicals	0.09	Utilities
0.17	Trade	0.08	Tobacco
0.16	Rubber and misc. plastics	0.06	Non-metallic mining
0.16	Fabricated metal	0.05	Metal mining
0.15	Machinery, non-electrical	0.01	Coal mining

Table D.3: Unweighted Two-Step Difference GMM Estimation Results without Lagged Dependent Variable

Dependent Variable	**Internal**	**External**	**All**
	η_b	η_b	η_b
Native Employment	-3.93**(1.97)	-9.75***(3.24)	-7.21***(2.02)
Hansen Test, p-value	0.64	0.36	0.15
$\mathrm{M}_1, p-val$	0	0.01	0
$\mathrm{M}_2, p-val$	0	0.01	0.01
GDP per Worker	1.91(2.49)	2.63*(1.35)	2.24*(1.14)
Hansen Test, p-value	0.18	0.42	0.23
$\mathrm{M}_1, p-val$	0.12	0.12	0.12
$\mathrm{M}_2, p-val$	0.05	0.06	0.06
Capital-Output	-0.06(0.08)	-0.01(0.11)	-0.07(0.07)
Hansen Test, p-value	0.69	0.56	0.83
$\mathrm{M}_1, p-val$	0.41	0.38	0.39
$\mathrm{M}_2, p-val$	0.35	0.34	0.3
Average Hours	0.25*(0.13)	0.23*(0.13)	0.29***(0.09)
Hansen Test, p-value	0.57	0.45	0.77
$\mathrm{M}_1, p-val$	0.85	0.79	0.88
$\mathrm{M}_2, p-val$	0.43	0.44	0.41
TFP	2.23(2.24)	2.21*(1.31)	2.08*(1.14)
Hansen Test, p-value	0.17	0.62	0.65
$\mathrm{M}_1, p-val$	0.14	0.14	0.15
$\mathrm{M}_2, p-val$	0.03	0.03	0.03
Skill Intensity	-0.08(0.21)	0.09(0.14)	0.02(0.12)
Hansen Test, p-value	0.33	0.15	0.14
$\mathrm{M}_1, p-val$	0	0	0
$\mathrm{M}_2, p-val$	0.04	0.03	0.03
Skill Bias	0.7(0.67)	2.48***(0.93)	2.03***(0.65)
Hansen Test, p-value	0.06	0.25	0.11
$\mathrm{M}_1, p-val$	0.26	0.47	0.38
$\mathrm{M}_2, p-val$	0.26	0.23	0.25
Hours of High-skilled	2.82***(0.99)	2.41***(0.91)	2.59***(0.79)
Hansen Test, p-value	0.27	0.32	0.24
$\mathrm{M}_1, p-val$	0.05	0.03	0.03
$\mathrm{M}_2, p-val$	0.6	0.6	0.62

Explanatory variables are immigrants as a share of total employment with a coefficient of η_b. Each cell is the result of a separate regression. The units of observations are 27 U.S. industries in each decade over $1960 - 2000$ and in 2005. Each regression includes year fixed effects. The method of estimation is 2-Step Difference GMM with Internal, External, and All instruments. Internal instruments are $(t-2; t-3)$ lags of immigrants' share in total employment. External instrument is an imputed share of immigrants in total employment based on the previous distribution of immigrants across industries. All instruments include both Internal and External instruments. The numbers in parentheses are heteroskedasticity robust and clustered by industry standard errors of the coefficients and (*) indicates significance level at 10, (**) at 5, and (***) at 1 percent. M_1 and M_2 are Arellano-Bond tests respectively on the first and second order serial correlation on the residuals.

Table D.4: Weighted Two-Step Difference GMM Estimation Results without Lagged Dependent Variable

Dependent Variable	**Internal**	**External**	**All**
	η_b	η_b	η_b
Native Employment	-5.62***(2.02)	-14.57***(3.86)	-11.33***(2.61)
Hansen Test, p-value	0.31	0.17	0.17
$\mathrm{M}_1, p-val$	0.01	0.04	0.02
$\mathrm{M}_2, p-val$	0.01	0.01	0.02
GDP per Worker	3.92(2.53)	5.84***(1.98)	5.85***(2.18)
Hansen Test, p-value	0.85	0.61	0.11
$\mathrm{M}_1, p-val$	0.08	0.07	0.08
$\mathrm{M}_2, p-val$	0.05	0.05	0.06
Capital-Output	-0.01(0.14)	0.17(0.18)	0.12(0.14)
Hansen Test, p-value	0.79	0.28	0.72
$\mathrm{M}_1, p-val$	0.83	0.72	0.77
$\mathrm{M}_2, p-val$	0.68	0.79	0.75
Average Hours	0.27**(0.12)	0.11(0.22)	0.13(0.19)
Hansen Test, p-value	0.31	0.26	0.55
$\mathrm{M}_1, p-val$	0.54	0.64	0.66
$\mathrm{M}_2, p-val$	0.67	0.96	0.93
TFP	4.12(3.01)	5.53**(2.59)	5.24**(2.28)
Hansen Test, p-value	0.86	0.46	0.11
$\mathrm{M}_1, p-val$	0.09	0.08	0.09
$\mathrm{M}_2, p-val$	0.06	0.06	0.07
Skill Intensity	-0.15(0.3)	0.41(0.31)	0.23(0.2)
Hansen Test, p-value	0.13	0.11	0.11
$\mathrm{M}_1, p-val$	0.01	0	0
$\mathrm{M}_2, p-val$	0.08	0.14	0.06
Skill Bias	0.01(1.67)	4.03***(1.5)	3.47***(1.22)
Hansen Test, p-value	0.06	0.13	0.24
$\mathrm{M}_1, p-val$	0.19	0.16	0.18
$\mathrm{M}_2, p-val$	0.33	0.57	0.53
Hours of High-skilled	1.69(1.08)	3.42***(1.32)	2.7***(0.94)
Hansen Test, p-value	0.22	0.46	0.54
$\mathrm{M}_1, p-val$	0.11	0.08	0.1
$\mathrm{M}_2, p-val$	0.39	0.3	0.23

Explanatory variables are immigrants as a share of total employment with a coefficient of η_b. Each cell is the result of a separate regression. The units of observations are 27 U.S. industries in each decade over $1960 - 2000$ and in 2005. Each regression includes year fixed effects and is weighted by the time-invariant average employment of industries over the period of $1960 - 2005$. The method of estimation is 2-Step Difference GMM with Internal, External, and All instruments. Internal instruments are $(t-2; t-3)$ lags of immigrants' share in total employment. External instrument is an imputed share of immigrants in total employment based on the previous distribution of immigrants across industries. All instruments include both Internal and External instruments. The numbers in parentheses are heteroskedasticity robust and clustered by industry standard errors of the coefficients and (*) indicates significance level at 10, (**) at 5, and (***) at 1 percent. M_1 and M_2 are Arellano-Bond tests respectively on the first and second order serial correlation on the residuals.

Table D.5: Unweighted Two-Step Difference GMM Estimation Results with Lagged Dependent Variable

Dependent Variable	Internal		External		All		All with Interaction Term		
	η_b	α_b	η_b	α_b	η_b	α_b	η_b	α_b	β_b
Native Employment	-2.59***(0.9)	0.85***(0.1)	-1.14(1.53)	0.9***(0.1)	-2.63***(1)	0.81***(0.08)	-2.73**(0.87)	0.77***(0.08)	0.02(0.02)
Hansen Test, p-value	0.17		0.36		0.22		0.33		
$M_1, p-val$	0.01		0.01		0.01		0.02		
$M_2, p-val$	0.33		0.49		0.37		0.18		
GDP per Worker	1.71(1.13)	0.88***(0.11)	1.17(1.53)	0.94***(0.17)	1.81*(1.01)	0.8***(0.1)	1.79(1.04)	0.78***(0.11)	-0.01(0.12)
Hansen Test, p-value	0.48		0.14		0.32		0.28		
$M_1, p-val$	0		0		0.01		0.01		
$M_2, p-val$	0.06		0.08		0.1		0.11		
Capital-Output	-0.01(0.07)	0.08(0.21)	0.05(0.08)	0.08(0.42)	-0.03(0.08)	0.07(0.28)	-0.03(0.06)	0.05(0.2)	0.01(0.05)
Hansen Test, p-value	0.87		0.74		0.81		0.74		
$M_1, p-val$	0.94		0.96		0.94		0.86		
$M_2, p-val$	0.81		0.86		0.75		0.73		
Average Hours	0.21*(0.11)	0.34**(0.16)	0.12(0.15)	0.42*(0.21)	0.18(0.14)	0.44*(0.21)	0.19*(0.08)	0.52**(0.16)	0.13***(0.03)
Hansen Test, p-value	0.39		0.32		0.42		0.65		
$M_1, p-val$	0.17		0.21		0.17		0.07		
$M_2, p-val$	0.09		0.12		0.08		0.12		
TFP	1.48(1.27)	0.93***(0.1)	0.88(1.63)	1.02***(0.13)	1.67(1.2)	0.87***(0.07)	1.6(1.2)	0.86***(0.07)	-0.01(0.14)
Hansen Test, p-value	0.53		0.15		0.29		0.28		
$M_1, p-val$	0		0		0		0		
$M_2, p-val$	0.04		0.04		0.06		0.06		
Skill Intensity	-0.06(0.05)	0.86***(0.09)	-0.14(0.11)	0.88***(0.11)	-0.07(0.06)	0.81***(0.09)	-0.05(0.06)	0.79***(0.11)	0.04(0.1)
Hansen Test, p-value	0.19		0.07		0.18		0.17		
$M_1, p-val$	0.02		0.02		0.02		0.06		
$M_2, p-val$	0.29		0.45		0.36		0.28		
Skill Bias	0.55(0.34)	0.59***(0.07)	1.2*(0.73)	0.54***(0.09)	0.51*(0.28)	0.59***(0.06)	0.47(0.31)	0.53***(0.07)	-0.1**(0.03)
Hansen Test, p-value	0.16		0.09		0.29		0.35		
$M_1, p-val$	0		0		0		0.01		
$M_2, p-val$	0.38		0.49		0.36		0.22		
High-Skilled Hours	1.25**(0.61)	0.33**(0.15)	1.65***(0.61)	0.2(0.14)	1.33***(0.42)	0.4***(0.12)	1.15**(0.4)	0.59***(0.1)	0(0.03)
Hansen Test, p-value	0.13		0.14		0.31		0.15		
$M_1, p-val$	0.77		0.65		0.45		0.06		
$M_2, p-val$	0.03		0.08		0.01		0		

Explanatory variables are immigrants as a share of total employment with a coefficient of η_b and $(t-1)$ lag of the dependent variable with a coefficient of α_b. Each cell is the result of a separate regression. The units of observations are 27 U.S. industries in each decade over $1960-2000$ and in 2005. Each regression includes year fixed effects. The method of estimation is 2-Step Difference GMM with Internal, External, and All instruments. Internal instruments are $(t-2; t-3)$ lags of both immigrants' share in total employment and dependent variable. External instruments are an imputed share of immigrants in total employment based on the previous distribution of immigrants. All instruments include both Internal and External instruments and "All with interaction term" adds to "All" regression an interaction of a dummy variable for year 2005 and lagged dependent variable with a coefficient of β_b. The numbers in parentheses are heteroskedasticity robust and clustered by industry standard errors of the coefficients and (*) indicates significance level at 10, (**) at 5, and (***) at 1 percent. M_1 and M_2 are Arellano-Bond tests respectively on the first and second order serial correlation on the residuals. The estimations use $(t-3)$ lags for internal instruments if there is a second-order serial correlation on the residuals as indicated by M_2.

Table D.6: Weighted Two-Step Difference GMM Estimation Results with Lagged Dependent Variable

Dependent Variable	Internal		External		All		All with Interaction Term		
	η_b	α_b	η_b	α_b	η_b	α_b	η_b	α_b	β_b
Native Employment	-1.93(1.36)	0.93***(0.08)	1.44(3.34)	1.02***(0.1)	-1.56(1.58)	0.89***(0.07)	-1.54(1.6)	0.92***(0.08)	-0.01(0.02)
Hansen Test, p-value	0.08		0.11		0.12		0.11		
$\mathrm{M}_1, p-val$	0.01		0		0.01		0.01		
$\mathrm{M}_2, p-val$	0.65		0.83		0.61		0.89		
GDP per Worker	1.58(1.54)	1.01***(0.14)	-0.5(1.55)	1***(0.1)	1.62(1.13)	0.62***(0.19)	1.98(1.26)	0.51**(0.26)	-0.11(0.08)
Hansen Test, p-value	0.77		0.28		0.38		0.18		
$\mathrm{M}_1, p-val$	0.01		0		0.11		0.49		
$\mathrm{M}_2, p-val$	0.04		0.06		0.03		0.04		
Capital-Output	-0.05(0.11)	0.07(0.24)	0.09(0.12)	0.11(0.26)	0.23*(0.12)	0.61(0.39)	0.25*(0.13)	0.63(0.4)	0.02(0.04)
Hansen Test, p-value	0.8		0.62		0.77		0.72		
$\mathrm{M}_1, p-val$	0.95		0.86		0.31		0.29		
$\mathrm{M}_2, p-val$	0.79		0.9		0.68		0.7		
Average Hours	0.03(0.07)	0.76***(0.11)	0.06(0.05)	0.74***(0.12)	0.71***(0.11)	0.01(0.07)	0.01(0.05)	0.963***(0.13)	0.12***(0.03)
Hansen Test, p-value	0.16		0.06		0.16		0.49		
$\mathrm{M}_1, p-val$	0.09		0.1		0.09		0.09		
$\mathrm{M}_2, p-val$	0.12		0.12		0.12		0.16		
TFP	1.38(1.71)	1.07***(0.11)	-1.21(1.39)	1.06***(0.17)	0.23(0.93)	0.85***(0.14)	0.97(0.83)	0.67***(0.2)	-0.21**(0.08)
Hansen Test, p-value	0.81		0.09		0.19		0.21		
$\mathrm{M}_1, p-val$	0.01		0.01		0.01		0.14		
$\mathrm{M}_2, p-val$	0.03		0.03		0.02		0.02		
Skill Intensity	-0.12(0.1)	0.59***(0.1)	-0.07(0.14)	0.59***(0.11)	-0.11(0.09)	0.67***(0.07)	-0.07(0.08)	0.68***(0.06)	-0.06*(0.03)
Hansen Test, p-value	0.1		0.07		0.21		0.22		
$\mathrm{M}_1, p-val$	0.25		0.21		0.15		0.01		
$\mathrm{M}_2, p-val$	0.21		0.11		0.09		0.12		
Skill Bias	-0.51(0.88)	0.68***(0.05)	-0.14(0.67)	0.66***(0.04)	-0.63(0.88)	0.68***(0.04)	-0.5(0.78)	0.68***(0.04)	-0.01(0.03)
Hansen Test, p-value	0.19		0.1		0.21		0.25		
$\mathrm{M}_1, p-val$	0.02		0.03		0.02		0.01		
$\mathrm{M}_2, p-val$	0.12		0.12		0.13		0.18		
High-Skilled Hours	0.25(0.95)	0.62***(0.11)	0.23(1.12)	0.61***(0.16)	0.35(0.87)	0.61***(0.12)	0.46(0.82)	0.61***(0.09)	0.02(0.02)
Hansen Test, p-value	0.25		0.13		0.51		0.32		
$\mathrm{M}_1, p-val$	0.16		0.2		0.17		0.1		
$\mathrm{M}_2, p-val$	0.07		0.07		0.06		0.6		

Explanatory variables are immigrants as a share of total employment with a coefficient of η_b and $(t-1)$ lag of the dependent variable with a coefficient of α_b. Each cell is the result of a separate regression. The units of observations are 27 U.S. industries in each decade over $1960-2000$ and in 2005. Each regression includes year fixed effects and is weighted by the total employment of the industry. The method of estimation is 2-Step Difference GMM with Internal, External, and All instruments. Internal instruments are $(t-2; t-3)$ lags of both immigrants' share in total employment and dependent variable. External instruments are an imputed share of immigrants in total employment based on the previous distribution of immigrants across industries. All instruments include both Internal and External instruments and "All with interaction term" adds to "All" regression an interaction of a dummy variable for year 2005 and lagged dependent variable with a coefficient of β_b. The numbers in parentheses are heteroskedasticity robust and clustered by industry standard errors of the coefficients and (*) indicates significance level at 10, (**) at 5, and (***) at 1 percent. M_1 and M_2 are Arellano-Bond tests respectively on the first and second order serial correlation on the residuals. The estimations use $(t-3)$ lags for internal instruments if there is a second-order serial correlation on the residuals as indicated by M_2.

Table D.7: Unweighted Two-Step Difference GMM Estimation Results with Lagged Dependent Variable varying by Elasticity of Substitution (σ) between High-skill and Low-skill Labor

Dependent Variable	**Internal**	**External**	**All**	**All with Interaction Term**	
	η_b	η_b	η_b	η_b	β_b
TFP: $\sigma = 1.75$	1.48(1.27)	0.88(1.63)	1.67(1.2)	1.6(1.2)	-0.01(0.14)
Hansen Test, p-value	0.53	0.15	0.29	0.28	
$\mathrm{M}_1, p-val$	0	0	0	0	
$\mathrm{M}_2, p-val$	0.04	0.04	0.06	0.06	
TFP: $\sigma = 1.5$	1.5(1.46)	1.21(1.72)	1.78(1.24)	1.63(1.24)	-0.01(0.15)
Hansen Test, p-value	0.62	0.16	0.29	0.27	
$\mathrm{M}_1, p-val$	0	0	0	0	
$\mathrm{M}_2, p-val$	0.06	0.07	0.08	0.09	
TFP: $\sigma = 2$	1.33(2.08)	0.72(1.84)	1.59(1.23)	1.54(1.23)	-0.02(0.14)
Hansen Test, p-value	0.5	0.06	0.28	0.27	
$\mathrm{M}_1, p-val$	0.01	0.01	0.01	0.01	
$\mathrm{M}_2, p-val$	0.05	0.03	0.06	0.06	
TFP: $\sigma = 5$	1.28(1.59)	0.42(1.8)	1.32(1.25)	1.27(1.23)	-0.03(0.15)
Hansen Test, p-value	0.41	0.05	0.26	0.23	
$\mathrm{M}_1, p-val$	0.01	0	0.01	0.01	
$\mathrm{M}_2, p-val$	0.51	0.03	0.06	0.06	
Skill Intensity: $\sigma = 1.75$	-0.06(0.05)	-0.14(0.11)	-0.07(0.06)	-0.05(0.06)	0.04(0.1)
Hansen Test, p-value	0.19	0.07	0.18	0.17	
$\mathrm{M}_1, p-val$	0.02	0.02	0.02	0.06	
$\mathrm{M}_2, p-val$	0.29	0.45	0.36	0.28	
Skill Intensity: $\sigma = 1.5$	-0.01(0.23)	-0.51(0.36)	-0.17(0.22)	-0.17(0.21)	-0.01(0.07)
Hansen Test, p-value	0.13	0.04	0.12	0.16	
$\mathrm{M}_1, p-val$	0	0	0	0	
$\mathrm{M}_2, p-val$	0.61	0.69	0.62	0.67	
Skill Intensity: $\sigma = 2$	-0.07(0.06)	-0.1(0.08)	-0.12(0.07)	-0.04(0.05)	0.27*(0.13)
Hansen Test, p-value	0.19	0.12	0.31	0.18	
$\mathrm{M}_1, p-val$	0.14	0.09	0.19	0.06	
$\mathrm{M}_2, p-val$	0.24	0.23	0.21	0.27	
Skill Intensity: $\sigma = 5$	0.29**(0.12)	0.12(0.19)	0.23**(0.11)	0.4***(0.09)	0.22***(0.04)
Hansen Test, p-value	0.06	0.03	0.19	0.41	
$\mathrm{M}_1, p-val$	0.45	0.24	0.13	0.14	
$\mathrm{M}_2, p-val$	0.41	0.39	0.49	0.45	

Explanatory variables are immigrants as a share of total employment with a coefficient of η_b and $(t-1)$ lag of the dependent variable. Each cell is the result of a separate regression. The units of observations are 27 U.S. industries in each decade over $1960-2000$ and in 2005. Each regression includes year fixed effects and is weighted by the total employment of the industry. The method of estimation is 2-Step Difference GMM with Internal, External, and All instruments. Internal instruments are $(t-2; t-3)$ and further lags of both immigrants' share in total employment and dependent variable. External instruments are an imputed share of immigrants in total employment based on the previous distribution of immigrants across industries. All instruments include both Internal and External instruments and "All with interaction term" adds to "All" regression an interaction of a dummy variable for year 2005 and lagged dependent variable with a coefficient of β_b. The numbers in parentheses are heteroskedasticity robust and clustered by industry standard errors of the coefficients and (*) indicates significance level at 10, (**) at 5, and (***) at 1 percent. M_1 and M_2 are Arellano-Bond tests respectively on the first and second order serial correlation on the residuals. The estimations use $(t-3)$ lags for internal instruments if there is a second-order serial correlation on the residuals as indicated by M_2.

Table D.8: Unweighted Two-Step Difference GMM Estimation Results with Lagged Dependent Variable Controlling for Task Specialization

Dependent Variable	**Internal**		**External**		**All**		**All with Interaction Term**	
	η_b	Γ_b	η_b	Γ_b	η_b	Γ_b	η_b	Γ_b
TFP	1.48(1.27)		0.88(1.63)		1.67(1.2)		1.6(1.2)	
Hansen Test, p-value	0.53		0.15		0.29		0.28	
$\mathrm{M}_1, p-val$	0		0		0		0	
$\mathrm{M}_2, p-val$	0.04		0.04		0.06		0.06	
TFP with task specialization	1.57(2.32)	-0.18(0.32)	0.7(1.66)	-0.26(0.18)	1.91(1.25)	-0.21(0.19)	1.77(1.18)	-0.22(0.17)
Hansen Test, p-value	0.36		0.06		0.21		0.18	
$\mathrm{M}_1, p-val$	0.01		0.01		0		0	
$\mathrm{M}_2, p-val$	0.05		0.03		0.06		0.05	
Skill Bias	0.55(0.34)		1.2*(0.73)		0.51*(0.28)		0.47(0.31)	
Hansen Test, p-value	0.16		0.09		0.29		0.35	
$\mathrm{M}_1, p-val$	0		0		0		0.01	
$\mathrm{M}_2, p-val$	0.38		0.49		0.36		0.22	
Skill Bias with task specialization	0.03(0.4)	0.15(0.09)	1.07(1.1)	0.1(0.14)	0.08(0.42)	0.16*(0.09)	0.03(0.08)	-0.04***(0.01)
Hansen Test, p-value	0.2		0.06		0.25		0.26	
$\mathrm{M}_1, p-val$	0		0		0		0.01	
$\mathrm{M}_2, p-val$	0.39		0.63		0.45		0.95	

Explanatory variables are immigrants as a share of total employment with a coefficient of η_b, $(t-1)$ lag of the dependent variable and communication-manual specialization of natives with a coefficient of Γ_b. Each cell is the result of a separate regression. The units of observations are 27 U.S. industries in each decade over $1960-2000$ and in 2005. Each regression includes year fixed effects and is weighted by the total employment of the industry. The method of estimation is 2-Step Difference GMM with Internal, External, and All Instruments. Internal instruments are $(t-2; t-3)$ lags of both immigrants' share in total employment and dependent variable. External instruments are an imputed share of immigrants in total employment based on the previous distribution of immigrants across industries. All instruments include both Internal and External instruments and "All with interaction term" adds to "All" regression an interaction of a dummy variable for year 2005 and lagged dependent variable. The numbers in parentheses are heteroskedasticity robust and clustered by industry standard errors of the coefficients and (*) indicates significance level at 10, (**) at 5, and (***) at 1 percent. M_1 and M_2 are Arellano-Bond tests respectively on the first and second order serial correlation on the residuals. The estimations use $(t-3)$ lags for internal instruments if there is a second-order serial correlation on the residuals as indicated by M_2.

Figure D.1: Immigration Trends in the U.S., 1990-2000 (Source: Singer, 2004)

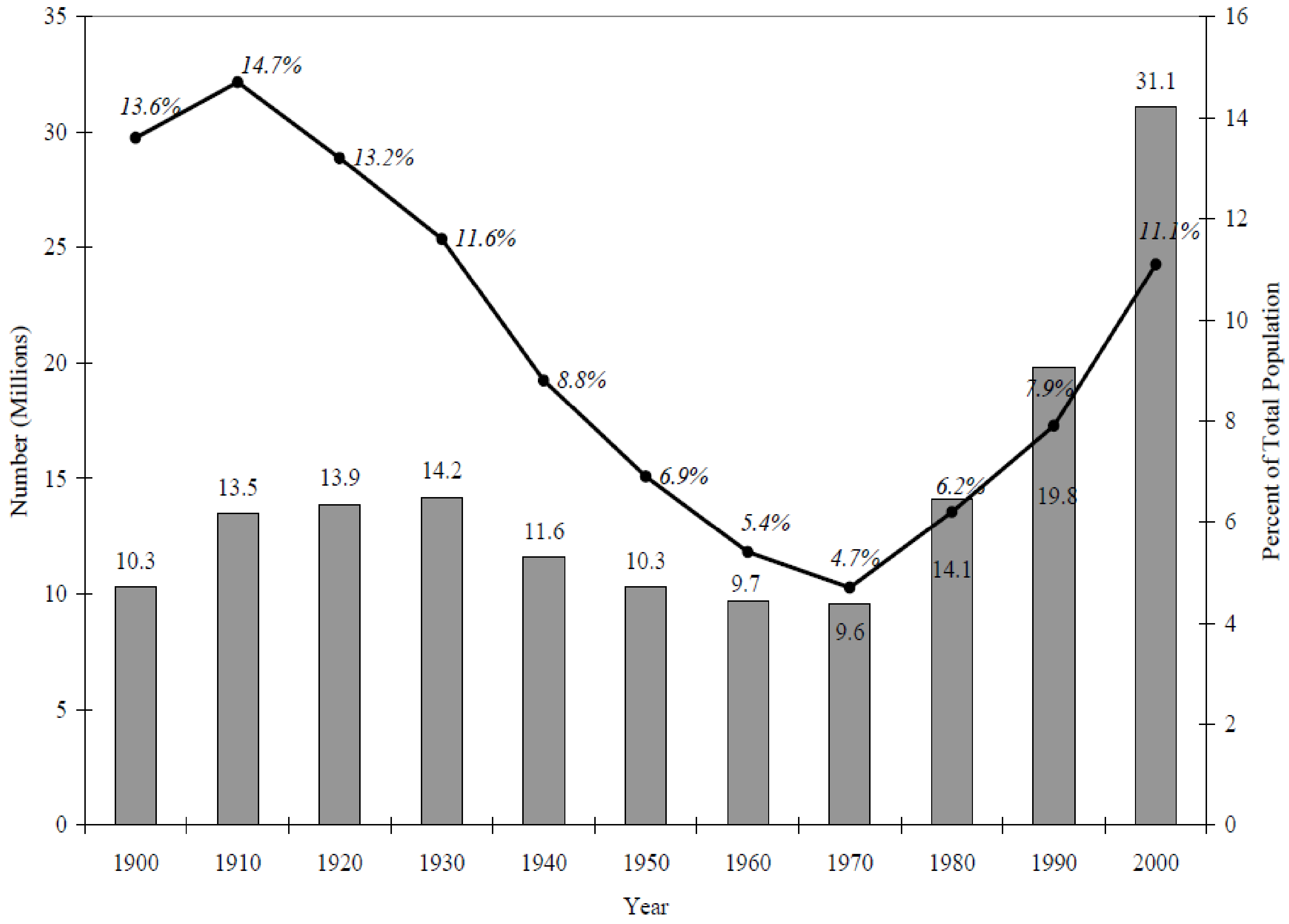


Figure D.2: Immigrants’ Distribution across U.S. States, 2006 (Source: Peri, 2012)

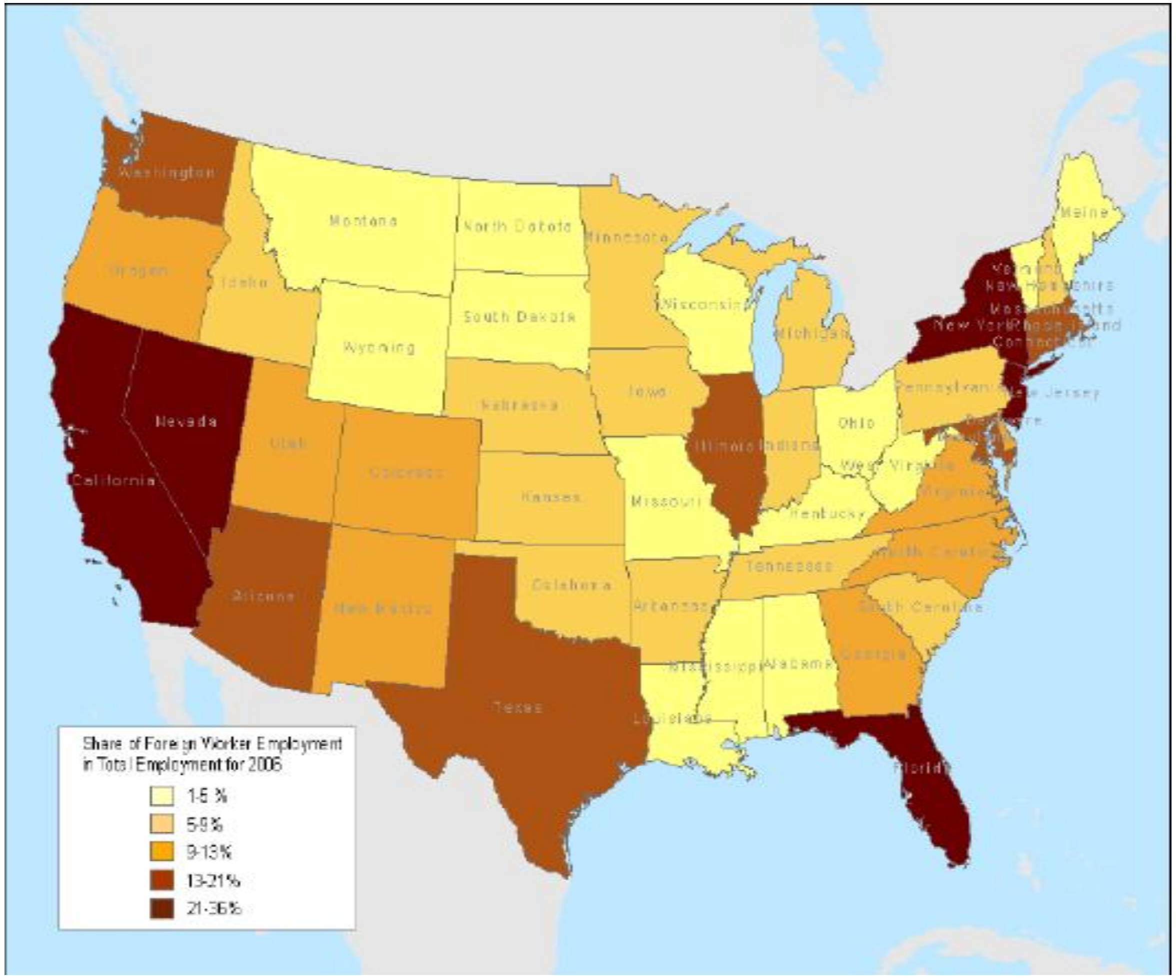


Figure D.3: Share of Immigrants in Employment by States and Industries relative to 1960, 1970-1980 (author's calculations based on U.S. Census data)

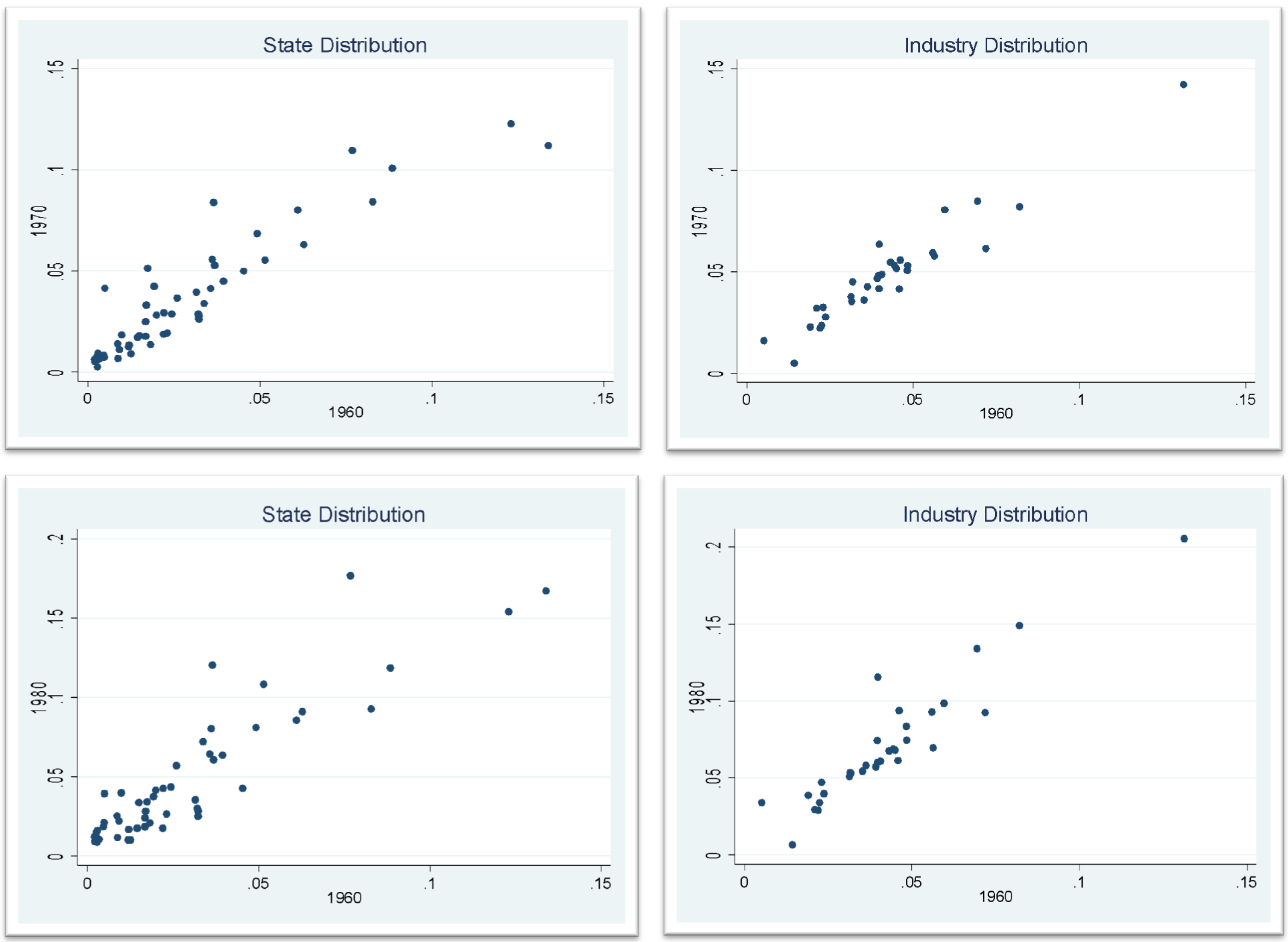


Figure D.4: Share of Immigrants in Employment by States and Industries relative to 1960, 1990-2000 (author's calculations based on U.S. Census data)